Niche partitioning in Great British bats through dietary specialisation

Roselyn Lydia Ware, BSc (Hons)

Submitted for the degree of Doctor of Philosophy

The University of Warwick
School of Life Sciences

July, 2016

Table of contents

List of figures

List of tables

Acknowledgements

Foremost, I would like to thank my supervisor Professor Robin Allaby. You have challenged and pushed me to pull off far more than I could have hoped to achieve. Your expertise, guidance, and forbearance, even in the face of my outbursts of melodrama, really kept me on the straight and narrow!

I owe a great deal of thanks to the members of the Allaby Lab past and present; Andrew Clarke, Rafal Gutaker, Hsiao-Lei Liu, Logan Kistler, James Kitchen, Ryan Morrison, William Nicholson, and Oliver Smith. I should also thank the wider community, both at Gibbet Hill, and at Wellesbourne, who are too numerous to mention (although, I would like to give a special mention to Hayley King for all of her support)!

This research was funded by a National Environment Research Council PhD studentship, with additional funding coming from Ecowarwicker Ecological Forensics. I also have to thank NERC for funding me to attend the NBAF Population Genomics Workshop in Liverpool, the CEH Keystone Skills in Bioinformatics in Wallingford, and the Statistics for Environmental Evaluation course in Glasgow.

This project would not have been possible without Ecowarwicker Ecological Forensics, which was the source of the lion's share of my samples. Additionally, I have depended greatly upon a number of individuals and organisations who have generously supplied me with guano samples for those species not well represented in the dataset. I especially thank David Bullock of the National Trust, who provided numerous samples from across the species. Jackie Underhill contributed the positive control sample; from a hand-reared *P. pygmaeus*. Hurrah for bat carers! I would also like to thank Ian Tanner and Lois Browne for providing samples, and the Warwickshire Bat Group for their support and feedback.

Thanks to Gemma Kay and Jackie Chan for help with the MiSeq, and to Mark Pallen for the use of the MiSeq. My advisory panel Dr Yin Chen and Dr Dave Chandler have been invaluable in their feedback and insights into my work. Thank you also to Steve Rossiter for his guidance.

I want to thank the School of Life Sciences IT support for keeping our servers going in the face of the abuse we put them through. Thanks too to the other users of the servers, to whom I am sure I have been a real pain.

Finally, on a personal note, I would like to thank my wonderful parents, Anne and Julian, my lovely sister Ellie, and the rest of my extended family. You may not always know what it is I *do*, but you have always encouraged me, and been the best twitter support team I could have hoped for! Last but not least, I want to thank Oli. Without your support would never even have considered applying for this PhD; so it's all your fault! You have been a rock!

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work was performed at the University of Warwick in the School of Life sciences. It was supervised by Professor Robin G. Allaby during the period of October 2012 to April 2016.

Some of the guano morphology data was generated by Benjamin Garrod in 2013 as part of his BSc dissertation.

A large portion of the DNA extractions was done by Robin Allaby, Rafal Gutaker, Oliver Smith, Ewa Samorzewska, Rabiah Saleh, and Daniel Carlotta-Jones as part of the Ecowarwicker Ecological forensics. This was done during the period of 2009-2015.

This work was presented in poster form at a number of conferences. Oral presentations of this work were given at the European Congress on Mammology in Stockholm, August 2015, at the National Bat Conference in Warwick, September 2015, and at the Britbats 3 Conference in London, April 2016.

This work has not yet been published in a peer-reviewed scientific journal.

Preface

Bats are vital to Great British biodiversity; they are the primary consumers of nocturnal insects, disperse nutrients across landscapes, and are excellent bioindicators of an ecosystem's health. The diversity of bat species in the UK is thought to be as a result of extensive resource partitioning.

There are numerous methods used for studying bat diets, each with benefits and drawbacks. Past research has compared small numbers of species at a time, making inter-species comparisons difficult. Our large repository of bat guano samples, collected from around the UK, has allowed us to study the bat species under one methodological 'umbrella'.

This thesis is divided into 7 chapters. This first chapter gives a broad overview of the project, framing this research and provides an overview of the technologies available, and how their development has enabled environmental research on a scale, which, until recently, would have been unimaginable. The second chapter is a meta-analysis of the literature that pertains to bat diets. These data will be used to inform the design of primers in the barcoding stages of the project.

Next is a shotgun metagenomic analysis of a selection of guano samples from across the range of the UK species. This method provides information, not only about diet species, but also about the bat, viral, and bacterial DNA. Analyses of this data show that there are several dietary forms seen between the species. The fourth chapter is a targeted amplicons metagenome study of the mitochondrial COI barcode region from the arthropod species identified in the literature review, and from metagenomic data-set. This provides a greater resolution picture of the diet species present. Analyses of these data use phylogenetic intersection analysis to ensure the robustness of the taxonomic assignments in the face of the patchy databases available.

In the fifth chapter, I draw together the data gathered using the different approaches and presented in the previous chapters. I discuss the efficacy of the methods, and assess the role of resource partitioning in bat species co-existence. The sixth section will look at the appropriateness of using guano morphology as a diagnostic of species presence. Finally, in chapter seven, I summarise these data in the wider context of bat ecology, comment on the implications of the research for conservation, and discuss potential directions for the field in the light of this research.

List of common abbreviations

A	adenine
aDNA	ancient DNA
ADR	artificial duplicate reads
ATP	adenosine tri-phosphate
BCT	Bat Conservation Trust
BLAST	Basic Local Alignment Search Tool
bp	base pairs
B. barbastellus	*Barbastella barbastellus*, western barbastelle bat (Schreber, 1774)
C	cytosine
CO1	cytochrome c oxidase subunit 1
CTAB	cetyltrimethylammonium bromide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide 5'-triphosphate
dsDNA	double stranded DNA
dsRNA	double stranded RNA
eDNA	environmental DNA
EDTA	ethylenediaminetetraacetic acid
E. serotinus	*Eptesicus serotinus*, serotine bat (Schreber, 1774)
FASTA	FAST-All format
FASTQ	FAST-Quality format
g	gram
G	guanine
GATC	GATC biotech
Gb	giga base pairs
GB	gigabyte
h	hour
HDC	high duty call
HGP	human genome project
HPLC	high performance liquid chromatography
HTS	high-throughput sequencing

indel	insertion-deletion
IUPAC	International Union of Pure and Applied Chemistry
kb	kilobase
kbp	kilobase pairs
KB	kilobytes
L	litre
LDC	low duty call
M	Molar (mol/L)
Mb	megabase
Mbp	megabase pairs
MB	megabytes
MEGAN	MEtaGenome ANalyzer
min	minute
ML	maximum likelihood
mM	millimolar
mtDNA	mitochondrial DNA
M. alcathoe	*Myotis alcathoe,* alcathoe bat (von Helversen and Heller, 2001)
M. bechsteinii	*Myotis bechsteinii*, Bechstein's bat (Kuhl, 1817)
M. brandtii	*Myotis brandtii*, Brandt's bat (Evermann, 1845)
M. daubentonii	*Myotis daubentonii*, Daubenton's bat (Kuhl, 1817)
M. nattereri	*Myotis nattereri*, Natterer's bat (Kuhl, 1817)
M. myotis	*Myotis myotis,* greater mouse-eared bat (Borkhausen, 1797)
M. mystacinus	*Myotis mystacinus*, whiskered bat (Kuhl, 1817)
na	nucleic acid
NA	not applicable
NERC	National Environment Research Council
NCBI	National Centre for Biotechnology Information
NEB	New England Biolab
NGS	next generation sequencing
NT	National Trust
N. leisleri	*Nyctalus leisleri*, Leisler's bat (Kuhl, 1817)

N. noctula	*Nyctalus noctula*, common noctule bat (Schreber, 1774)
OTU	operational taxonomic unit
PCA	principal components analysis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIA	phylogenetic intersection analysis
pg	picogram
pM	picomolar
P. pipistrellus	*Pipistrellus pipistrellus*, common pipistrelle bat (Schreber, 1774)
P. pygmaeus	*Pipistrellus pygmaeus,* soprano pipistrelle bat (Leach, 1825)
P. nathusii	*Pipistrellus nathusii,* Nathusius' pipistrelle bat (Keyserling & Blasius, 1839)
P. auritus	*Plecotus auritus,* brown long-eared bat (Linnaeus, 1758)
P. austriacus	*Plecotus austriacus*, grey long-eared bat (Fischer, 1829)
R. hipposideros	*Rhinolophus hipposideros*, lesser horse-shoe bat (Bechstein, 1800)
R. ferrumequinum	*Rhinolophus ferrumequinum*, greater horse-shoe bat (Schreber, 1774)
RNA	ribonucleic acid
RPC	reverse-phase cartridge
rRNA	ribosomal RNA
s	second
SMRT	single molecular real-time sequencing
SNP	single nucleotide polymorphism
sp.	species (singular)
spp.	species (plural)
ssDNA	single stranded DNA
ssp.	subspecies

ssRNA	single stranded RNA
T	thymine
T_m	melting temperature
TAE	Tris acetate EDTA
TB	Terabyte
Tris-HCL	tris(hydroxymethyl)aminomethane hydrochloride
µg	microgram
µl	microliter
µM	micromolar
°C	degrees Celsius

Chapter 1 : Introduction

Predation of arthropods by bats is vital in maintaining the balance of the global ecosystem, and plays a crucial role in the control of insect pests, notably agricultural pests (López-Hoffman et al., 2014, Maine and Boyles, 2015). Additionally, the particular dietary choices of different bat species are thought to be fundamental in supporting their diversity. This thesis explores some of the methods available for measuring the diets of bats, evaluates the efficacy of each method, and attempts to characterise the diets of all of the Great British bat species. I determine and describe the dietary niches of each of the bats, and identify the overlap of these niches. Accurate information about the diets of bats allows inferences to be made about potential competition between species; the impacts of habitat destruction and climate change; and the potential extinction risk of the bats (Boyles and Storm, 2007, Kunz et al., 2011, Pineda-Munoz and Alroy, 2014, Safi and Kerth, 2004).

1.1. The global importance of bats and the threats they face

There are over 1100 known species of bats worldwide (Wilson and Reeder, 2005). After rodents, bats (order Chiroptera) are the second largest group of mammals. Bats are pivotal to supporting global biodiversity; they are the primary consumers of nocturnal insects, disperse nutrients across landscapes, and are excellent bioindicators of an ecosystem's health (Jones et al., 2009, Patterson et al., 2003).

However, 25% of bats worldwide are classed as 'of conservation concern', with a further 21% classified as 'near threatened' (Boyles and Storm, 2007). Four of the seventeen Great British species are listed as declining, with the status of several others unknown (I.U.C.N., 2013). The plethora of threats faced by bats include: unsympathetic development projects, destruction of tree lines and hedgerows, the drainage of wetlands, infectious diseases, and the impact of pesticides (Ashrafi et al., 2011, Mickleburgh et al., 2002). Additionally, climate changes may have a

highly detrimental impact on bats, including, but not limited to: changes in prey abundances, alterations in the efficacy of echolocation calls, and the consequences of extreme weather events (Luo et al., 2014).

1.2. Using diet to define the ecological niche

The ecological niche is, by no means, a recent concept (Grinnell, 1917), and the thorough understanding of the ecological niche of an organism is considered vital to its conservation (Ashrafi et al., 2011, Sattler et al., 2007). Gause's principle of competitive exclusion states that the partitioning of the ecological niche by the division of limiting resources is the fundamental mechanism by which species can stably co-exist, while avoiding extensive interspecific competition (Hardin, 1960, Schoener, 1974). Typically, the most important factors in bat niche separation are considered to be the partitioning of habitat and of diet (Ashrafi et al., 2011, Schoener, 1974). Where the niches of two species are highly similar (i.e. the niche overlap is great) out-competing can occur if the resource in question is limited.

The ranges of many of the Great British bat species are acutely overlapping (Bat Conservation Trust, 2009), suggesting that trophic resource partitioning is important in supporting the species in Great Britain (see appendix C.1.1) (Aguirre et al., 2002, Arlettaz et al., 2000). By forming a picture of how dietary arthropods are partitioned between bat species, it becomes possible to ascertain to the level of dietary niche overlap, which can inform conservation and management strategies. In order to identify dietary overlaps, a thorough knowledge of the bat diets is key. Knowledge of bat diets also gives information dietary breadth, which has been shown to be strongly correlated with extinction risk (Boyles and Storm, 2007, Safi and Kerth, 2004).

Information about bat diets benefits conservation efforts. It is, however, important to bear in mind that the conservation of bat species is not merely an altruistic act; bat predation of agricultural insect pests is a key

ecosystem service postulated to provide significant economic benefit (Kunz et al., 2011). By quantifying ecosystem services in economic terms in an understandable and easily transferable metric, it becomes far easier to convince policy makers of the importance of conservation (Costanza et al., 1997).

1.3. Factors that influence bat diets

A number of variables will impact the diet of a bat, aside from the species of that bat. An understanding of these factors, in combination with knowledge of the diets would contribute greatly to conservation efforts.

1.3.1. Feeding style

Fenton and Bogdanowicz identify four major feeding styles: aerial feeders (those that hunt airborne prey), gleaners (those that take prey directly from the ground or vegetation), bats which feed primarily over water (from flighted arthropods), and those which use trawling (taking arthropods directly from the water surface) (Fenton and Bogdanowicz, 2002). Table 1.1 shows the breakdown of the feeding styles of the Great British bats. Data is currently unavailable on *M. alcathoe*. Wing shape, body size, and tail length, all have impacts on flight and therefore on feeding styles (Fenton and Bogdanowicz, 2002, Norberg, 1995).

Table 1.1: Feeding styles and the emergence times of Great British bats
Adapted from (Fenton and Bogdanowicz, 2002). A- aerial hawking, G- gleaning, W- feeding over water, T- trawling, DD- data deficient. Sorted by emergence time (mins after sunset).

Species	Feeding Style	Feeding Style Reference	Emergence Time (min after sunset)	Emergence Time Reference
N. noctula	A	(Holderied and Von Helversen, 2003)	5	(Jones and Rydell, 1994)
P. nathusii	A	(Holderied and Von Helversen, 2003)	11-50	(Gelhaus and Zahn, 2010)

N. leisleri	A	(Holderied and Von Helversen, 2003)	18	(Jones and Rydell, 1994)
B. barbastellus	A	(Holderied and Von Helversen, 2003)	19.5±6.6	(Russo et al., 2007)
E. serotinus	A/W	(Catto et al., 1994, Catto et al., 1996, Holderied and Von Helversen, 2003, Jones, 1993, Robinson and Stebbings, 1997)	20	(Jones and Rydell, 1994)
P. pygmaeus	A	(Holderied and Von Helversen, 2003)	24.8	(Davidson-Watts and Jones, 2006)
R. ferrumequinum	A/G	(Ahmim and Moali, 2013, Fenton, 1997, Jones et al., 1995, Jones and Rayner, 1989)	25	(Jones and Rydell, 1994)
M. brandtii	A	(Fenton and Bogdanowicz, 2002)	30	(Middleton et al., 2014)
P. austriacus	A	(Bauerova, 1982, Razgour et al., 2011)	30	(Middleton et al., 2014)
R. hipposideros	A/G	(Ahmim and Moali, 2013, Jones and Rayner, 1989)	31	(Jones and Rydell, 1994)
M. mystacinus	A	(Fenton and Bogdanowicz, 2002)	32	(Jones and Rydell, 1994)

P. pipistrellus	A	(Holderied and Von Helversen, 2003)	32	(Jones and Rydell, 1994)
M. bechsteinii	A/G	(Fenton and Bogdanowicz, 2002, Petrov, 2006, Wolz, 1993)	33	(Jones and Rydell, 1994)
P. auritus	G	(Coles et al., 1989)	54	(Jones and Rydell, 1994)
M. nattereri	A/G	(Arlettaz et al., 1997, Geisler and Dietz, 1999, Jones, 1993, Shiel et al., 1991, Siemers and Schnitzler, 2000, Vaughan, 1997)	75	(Jones and Rydell, 1994)
M. daubentonii	T/W	(Boonman et al., 1998, Jones and Rayner, 1988, Kalko and Schnitzler, 1989, Rydell et al., 1994, Rydell et al., 1999)	84, 28.2±12	(Jones and Rydell, 1994, Lučan, 2009)
M. alcathoe	DD		DD	

1.3.2. Predation using echolocation

Echolocation calls are used as a navigation and spatial orientation aid, and as a tool for locating insect prey. Echolocation is used by around 80% of bat species and all Great British bats (Schnitzler et al., 2003, Walters et al., 2013). There are a broad range of different echolocation

call frequencies and patterns displayed by the different species (Bogdanowicz et al., 1999). Echolocation strategy is strongly linked with feeding strategy, diet, and habitat preferences (Neuweiler, 1990). This difference is particularly marked in the two rhinolophid bats *R. hipposideros* and *R. ferrumequinum* (Yinpterochiroptera) when compared to the other Great British bats (all of which fall in the Yangochiroptera), which is likely due to their possible separate origin of echolocation; based on evidence from the duty cycle. The duty cycle of a call is the ratio of the signal duration (time between the onset and offset of a sound) and the period (the time between the onset of the successive sounds) (Fenton et al., 2012, Ho et al., 2013). Rhinolophid bats (Yinpterochiroptera) use a high duty cycle (HDC) call; they call >50% of the time, separating the call pulse and echo information by varying the frequency of the call. Low duty cycle (LDC) calls separate pulse and echo information in time by waiting for the echo before issuing the next pulse, calling <20% of the time (Ho et al., 2013). LDC is used by the Yangochiroptera (Bogdanowicz et al., 1999, Fenton et al., 1995). Table 1.2 shows the echolocation call frequencies and styles most commonly used by the bat species. The origins of these two forms of echolocation calls (LDC and HDC), and the implications for taxonomy are discussed in section 1.6.2.

Table 1.2: Echolocation call frequencies and styles

Taken from (Stebbings, 1986), 1986. DD= data deficient. LDC= low duty cycle, HDC= high duty cycle.

Species	**Detected range of frequencies (kHz)**	**Pulse repetition rate in free flight (pulse/sec)**	**Description of sounds heard**	**Duty Cycle**	**Reference**
B. barbastellus	25-100	15-25	'Ticks'	LDC	(Stebbings, 1986)
E. serotinus	20-60	7-8	Hard 'tocks'	LDC	(Stebbings, 1986)
M. alcathoe	43-120	12	DD	LDC	(Von Helversen et

					al., 2001, Jan et al., 2010)
M. bechsteinii	20-100	20-30	'Ticks'	LDC	(Stebbings, 1986)
M. brandtii	35-105	10-15	'Clicks'	LDC	(Stebbings, 1986)
M. daubentonii	30-100	12-15	'Ticks'	LDC	(Stebbings, 1986)
M. mystacinus	35-105	10-15	'Clicks'	LDC	(Stebbings, 1986)
M. nattereri	30-100	20-35	'Ticks'	LDC	(Stebbings, 1986)
N. leisleri	15-80	8-10	Metallic 'chinks'	LDC	(Stebbings, 1986)
N. noctula	15-75	5-10	Metallic 'chinks'	LDC	(Stebbings, 1986)
P. nathusii	34-79	35-50	'Plops'	LDC	(Bat Conservation Trust, 2013, Russ and Racey, 2007)
P. pipistrellus	40-83	12-28	'Smacks'	LDC	(Bat Conservation Trust, 2013, Russ and Racey, 2007)
P. pygmaeus	47-90	10-20	'Smacks'	LDC	(Bat Conservation Trust, 2013, Russ and Racey, 2007)
P. auritus	25-100	15-25	'Ticks'	LDC	(Stebbings, 1986)
P. austriacus	25-100	15-25	'Ticks'	LDC	(Stebbings, 1986)
R. ferrumequinum	40-100	8-11	Soft warbles	HDC	(Stebbings, 1986)
R. hipposideros	60-125	10-15	Soft warbles	HDC	(Stebbings, 1986)

Ear size and shape can vary dramatically between the different bat species and are adapted to the echolocation style of the bat. Additionally, the rhinolophid bats (*R. hipposideros* and *R. ferrumequinum*) have distinctive nose leaves, which aid echolocation by enhancing the range and sensitivity of call directionality (Bogdanowicz et al., 1997).

There is a biological arms race between bats and their prey, many of which have adopted elaborate mechanisms to evade being preyed upon (Miller and Surlykke, 2001). Some flying insects have developed the ability to hear the ultrasonic calls made by bats, allowing them to avoid predation (Ter Hofstede et al., 2013, Yager, 2012). Saturniid moths have evolved spinning hind-wing tails, which entice the bats, and act as a decoy (Barber et al., 2015). Some species of moths have developed echolocation calls as a means of acoustic aposematism (Chesmore, 2004, Corcoran and Hristov, 2014), through Müllerian mimicry (Müller, 1878) (where harmful species mimic each other's warning signals) and Batesian mimicry (Bates, 1981) (whereby harmless species mimic the warning signals of distasteful or harmful species) (Barber and Conner, 2007). Whereas other species have developed methods of sonar jamming (Corcoran and Conner, 2012). However some species of bat have adapted to locate arthropods by the acoustic detection of echolocation calls generated by the prey (Falk et al., 2015). One such example is *Eptesicus fuscus,* which is able to detect the echolocation calls of the Tettigoniidae (bush crickets) (Neuweiler, 1990). These strategies may have differential success rates on the different bat species, potentially influencing predation, and consequentially, diet.

1.3.3. Dentition and feeding

Dentition is adapted to the bat dietary preferences, a pattern which is also reflected in other mammalian taxa (Gill et al., 2014). Insectivorous bats which feed on hard bodied prey such as Coleoptera (beetles) are thought to be characterised by powerful mandibles and robust crania, longer canines, and larger teeth than those which feed on soft bodied prey

(Freeman, 1979, Freeman, 1998, Ghazali and Dzeverin, 2013). Additionally, the size of the prey is a key determiner of predation likelihood by a given bat; prey that is too small is energetically inefficient to feed upon, whereas prey that is too large may be difficult to catch. Furthermore, large insects are often harder, and may exceed the bat's maximum bite size (Freeman and Lemen, 2007).

1.3.4. Emergence time and prey availability

Prey availability varies throughout the day and through the seasons. The time that bats emerge to feed (table 1.1) varies greatly between species, even between those hunting in the same locations. The crepuscular and nocturnal hunting of bats is hypothesised to be in order to prevent avian predation on bats and competition with bird species (Jones and Rydell, 1994). However, the crepuscular peak flight activity of many potential prey forces earlier emergence of some bats. Typically, the fast-flying aerial hawking bats emerge earlier, as they are better adapted for evading being preyed upon, whereas gleaners and those that feed predominantly on moths, emerge later (Jones and Rydell, 1994). This can be seen in table 1.1.

1.3.5. Distributions and habitat selection of Great British bats

All of the Great British bat species have ranges which include the south of England, with only around half have ranges extending to Scotland (*M. myotis, M. nattereri, M. daubentonii, N. noctula, N, leisleri, P. nathusius, P. pipistrellus,* and *P. auritus*) (Bat Conservation Trust, 2015). The ranges of the bats in Great Britain can be found in appendix B.1.1. As arthropod ranges vary across Great Britain, bat diets will also likely vary between different populations of the same species across the breadth of the country. The habitat occupied by a bat species will impact the prey species available, and thus which arthropods are likely to be in the diet. *M. myotis* (a species found across Europe) will select habitats suitable for accessing ground dwelling prey, such as mown grass, and forests without undergrowth, whereas the non-UK species *M. blythii* feeds on grass

associated prey, found in un-mown grasslands and pastures (Arlettaz, 1999). Many of the habitats relied upon by bats have been greatly affected by anthropogenic pressures. *P. pipistrellus* have been shown to exploit streetlights for feeding, this has led to population increases, which, as a result, is thought to have contributed to the decline of *R. hipposideros* populations due to the similarity in diets (Arlettaz et al., 2000). Furthermore, temperate bat ranges are predicted to be heavily impacted by climate change; under the most extreme climate forecasts, all temperate bat species in Europe (including all British species) are predicted to have contracted populations as a result of available ranges contracting (Rebelo et al., 2010).

1.3.6. Migration

A number of Great British bats are migratory. There is significant evidence that *P. nathusii* migrate between Great Britain and Europe, using offshore oil platforms as refuge points (Boshamer and Bekker, 2008, Lagerveld et al., 2014). Natural wetlands are crucial for supporting *P. nathusii* on their migration through the Iberian Peninsula; they will positively select natural wetlands, and avoid artificial wetlands such as paddy fields (Flaquer et al., 2009). Seasonal habitat selection will potentially result in variations in diet throughout the seasons.

1.4. Measuring the diets of bats

Direct observation of predation of insect in bats is very difficult (Clare et al., 2009). Traditional analysis of bat diets has relied heavily on microscopic analysis of digested insect fragments found in guano (Clare et al., 2011). However, bats thoroughly masticate and digest their prey, often discarding the hard to digest fragments such as the carapace or elytra (Bohmann et al., 2011, Rabinowitz and Tuttle, 1982, Zeale et al., 2011). This increases the likelihood of misidentification, and over representation of the tougher remains that were not discarded. Identifications made in this manner are rarely more specific than order level (Clare et al., 2009).

Other methods used to measure bat diets include the dissection of stomach contents, the identification of discarded prey remains, and stable isotope analysis. These methods are discussed in detail in chapter two.

1.4.1. Using environmental DNA sequencing

With the development of next generation 'high throughput' DNA sequencing methods, it has become possible to generate millions of DNA sequence reads in parallel, dramatically reducing time and costs. This has been widely exploited for studying environmental DNA (eDNA) from a variety of sources (Bohmann et al., 2011). Increasingly, molecular metagenomic techniques are used in analysing the diets of vertebrates.

1.4.2. Metagenomics

There are two main approaches used in metagenomics. The first is metabarcoding (sometimes referred to as amplicon metagenomics), PCR directed sequencing is used to target regions which are conserved enough across the organisms of interest to amplify using universal primers, but which are variable enough to allow (ideally) species level identification. Additionally, the target chosen often has a high copy number within each cell, such as is the case with plastid or mitochondrial targets. 16S ribosomal DNA (rDNA) amplicons have been widely used in the identification of a range of un-culturable prokaryotes (Wang and Qian, 2009). However, increasingly, a wider range of amplicon targets are being used, allowing eukaryotic targets to be studied alongside prokaryotic species. One example is cytochrome c oxidase subunit 1 (CO1) which is used by the Barcode Of Life Database (Ratnasingham and Hebert, 2007).

The metabarcoding approach can provide a high resolution picture of the species present in the sample, although there are a number of key limitations. First is the assumption that the primers used to amplify the target DNA will amplify only the appropriate DNA, and will amplify all of

the target species equally. This is particularly an issue where proportions of sequences classified to an organism is used as a proxy for the proportion of said organism's DNA in the sample, highlighting the requirement for good PCR primer design. The issue with the design of PCR primers is that they require some *a-priori* knowledge of the target organisms expected to be present in the sample; ideal primers will amplify all of the (for example) arthropod DNA from a sample, and nothing else, to maximise the efficiency of the sequencing. However, primers are designed based on knowledge of target sequences from sequence databases, and poorly studied taxa may be under represented in the databases. This means that these taxa can be missed during the primer design process, and as a result the primers may fail to amply DNA from these taxa despite their presence in the sample, leading to issues of bias. Alternatively, these primers may spuriously amplify non-target DNA. These issues can be further confounded by stochastic errors introduced during the PCR enrichment, whereby some PCR templates may be preferentially amplified over others (Best et al., 2015). The reference DNA sequence database (such as the NCBI's nt database (Altschul et al., 1990)) used to assign taxonomic classifications to the output sequences can have large impacts on the success of taxonomic classifications; using a patchy database in which many taxa are not represented will not only increase the chance of sequences being unassigned, but will increase the likelihood of sequences being misassigned (Smith et al., 2015). Furthermore, where poor quality DNA is used, fragmentation of the target region can cause PCR failure, resulting in a type II error.

The second major approach used is shotgun metagenomics. This is the sequencing of DNA in a sample without selecting for a particular target region. As a result, the endogenous DNA of the organisms of interest, such as bat prey, can be low in comparison to other DNA sources, depending on the type of sample used. Additionally, the database coverage is typically far more restricted than those for amplicon targets for any one genomic region. This causes a greater risk of

misassignments of sequences to over-represented (e.g. model organism) sequences in the database. However, without the need for the *a-priori* knowledge required for primer design, some ascertainment biases are avoided. Additionally, due to the use of fewer rounds of PCR in the sample preparation process used for preparing shotgun metagenome, there is less PCR stochastic bias. Shotgun metagenomics can also be used to provide metadata for the target data, such as gut microbiome data when studying diet. Direct sequencing allows identification of organisms which cannot be cultured, or cannot be distinguished from other species by using targeted sequencing such as 16S rDNA (Tringe et al., 2005). A major advantage of this approach is that it exploits more of the DNA laid down by the organisms than barcodes, and so has the potential to be more sensitive than metabarcoding especially in degraded samples (Smith et al., 2015). This is particularly true where the barcodes used cannot distinguish between taxa at a species level due to ambiguities in the sequences (Srivathsan et al., 2015). The development of methods that account for patchy database representation is of key importance when working with metagenomic data, and will be discussed at greater length in chapter three.

1.4.3. Applying faecal metagenomics to study mammalian diets

Metagenomics using faeces often focuses on the microbiome (Handelsman, 2004, Riesenfeld et al., 2004). Using many standard library preparation methods (such as Illumina TruSeq) bacterial DNA can easily be studied. Of the viruses, only dsDNA can be prepared without the further processing required for ssDNA and RNA viruses.

Increasingly faecal metagenomics is being used in order to study eukaryotic DNA to measure the diets of mammals, birds and reptiles (Jedlicka et al., 2013, Deagle et al., 2005, Tollit et al., 2009, Pompanon et al., 2012). A number of bat diets have also been characterised in this way, demonstrating the applicability of faecal metagenomic methods for studying bat diets (Bohmann et al., 2011, Razgour et al., 2011, Zeale et

al., 2011, Clare et al., 2009) amongst a number of other studies. However, to date, all studies have taken a metabarcoding approach, many of which used the primer set designed by Zeale et al. (2011). Furthermore, the diets of all of the different bat species present in Great Britain have not been comprehensively studied. Here we employ both metagenomic (metabarcoding and shotgun metagenomics) approaches to characterise the diets of Great British bats in an unbiased manner.

1.4.4. Technical considerations when working with DNA from guano

DNA from guano is typically of low quality; it may have a very low concentration, and, due to the degradation of DNA post cell-death, is often highly fragmented (Deagle et al., 2006). This can lead to issues related to contamination, allelic dropout, false alleles, or to the failure of the PCR primers to anneal or extend on the target DNA (Pompanon et al., 2005, Puechmaille et al., 2007). Additionally, there may be exogenous and endogenous nucleases, as well as components of bat guano, such as bile salts, complex polysaccharides, and urea, which may act as PCR inhibitors (Idaghdour et al., 2003, Khan et al., 1991, Lantz et al., 1997, Monteiro et al., 1997).

The field of aDNA research has been instrumental in developing methods to avoid contamination, and to deal with other technical challenges arising from inhibition, fragmentation and other damage such as cytosine deamination (Kistler et al., 2015, Pääbo et al., 2004, Smith et al., 2003). Fragmented DNA with low endogenous concentration and target copy number, such as the DNA from guano, is vulnerable to contamination. PCR contamination is likely to be the most problematic as it could give confounding or false positive result, particularly where PCR products are from potential target organisms. Consideration of these factors is crucial when working with DNA from guano and must be considered as the field of faecal metagenomics develops. Methods used to address these issues will be discussed in chapters three and four, and appendix A.1.

To address the problems of DNA fragmentation, primers used to identify the bat species, and those for arthropod barcoding, have short amplicons and amplify mitochondrial genes which have a high copy number which increases the likelihood of an intact target region, thus increasing the likelihood of successful amplification. Additionally, mitochondrial genes have a higher mutation rate than genes located within nuclear DNA, enhancing their power to discriminate between species. Furthermore, these PCR targets are well represented on the databases such as GenBank (Altschul et al., 1990, Benson et al., 2000, Ratnasingham and Hebert, 2007). Primers must be able to amplify the range of target species required. Where the target gene is not highly conserved across the target species, a mixture of degenerate primers (i.e. some positions on the oligonucleotide may have one or more of the possible bases represented) may be used.

1.5. Environmental context of the source material

The Allaby group runs an ecological forensics service using the approaches of ancient DNA research to identify animal species from guano samples. The samples that are sent in allow us to build up a picture of the genetic diversity and evolutionary history of the species involved. At the time of study, there were over 5000 bat guano samples in the collection, with all Great British bat species represented, as can be seen in figure 1.1.

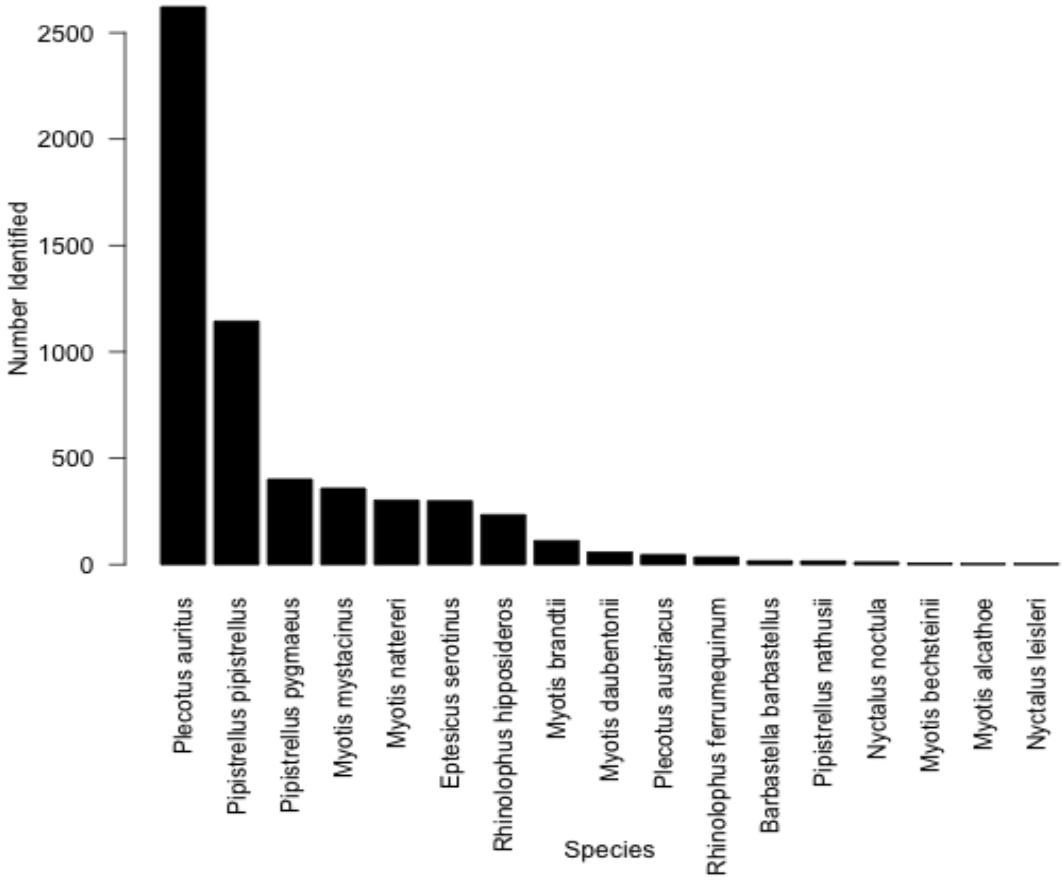


Figure 1.1: Bat species identified by the BatID service as of the beginning of this study

The samples accumulated cover the ranges of the bats from Great Britain. The samples used in this study have mostly been taken from the BatID service, as described in chapter three.

1.6. The importance of reliable bat identification

It is crucial when studying bat diets that the bat species from which the sample is taken is robustly identified. This is particularly key when studying bat species that are morphologically cryptic such as *P. pipistrellus* and *P. pygmaeus*, or the small myotis bats (*M. brandtii, M. daubentonii*, and *M. alcathoe*), or where guano samples are collected from the environment (as the guano is often morphologically cryptic, see chapter 6).

1.6.1. The classification of bats

The Chiroptera are traditionally divided into two suborders: the Megachiroptera, which are largely frugivorous and rarely use echolocation, and the Microchiroptera, which use echolocation and are mostly insectivorous. Microchiroptera are identifiable from the Megachiroptera by their small eyes, their short snouts, by the way their wings fold along their bodies, rather than wrap around, and often by distinctive facial features such as noseleaves, tragi, and large ears, which are adapted for echolocation. Size is, possibly surprisingly, not the most useful distinguisher, as the largest known Microchiroptera has a wing-span of nearly a meter; much larger than the smallest Megachiroptera, although the Microchiroptera are typically smaller (Richardson, 2011).

1.6.2. Updating the traditional taxonomy

The traditional classification, which is based primarily on morphological characteristics, implies monophyly of the group, with a single origin of both flight, and a single origin of echolocation (Simmons and Geisler, 1998). However, there is evidence that the rhinolophid Microchiroptera are, in fact, a sister taxon of the Megachiroptera, implying that either there was a single origin of echolocation in all bats, with the resultant loss of echolocation in the Megachiroptera, or that there were multiple origins of echolocation (Bates et al., 2001, Hill, 1974, Hoofer et al., 2003, Simmons, 1993, Van Den Bussche and Hoofer, 2004). There is still uncertainty as to the origins of echolocation, however, the two suborders Yinpterochiroptera/Pteropodiformes (previously the Megachiroptera, and Rhinolophoidea, previously of the Microchiroptera) and the Yangochiroptera/ Vespertilioniformes (the remaining Microchiroptera) are now generally accepted (figure 1.2) (Jones and Teeling, 2006, Teeling et al., 2005, Teeling, 2009, Li et al., 2008). The terms Yinpterochiroptera and Yangochiroptera will be used throughout this thesis.

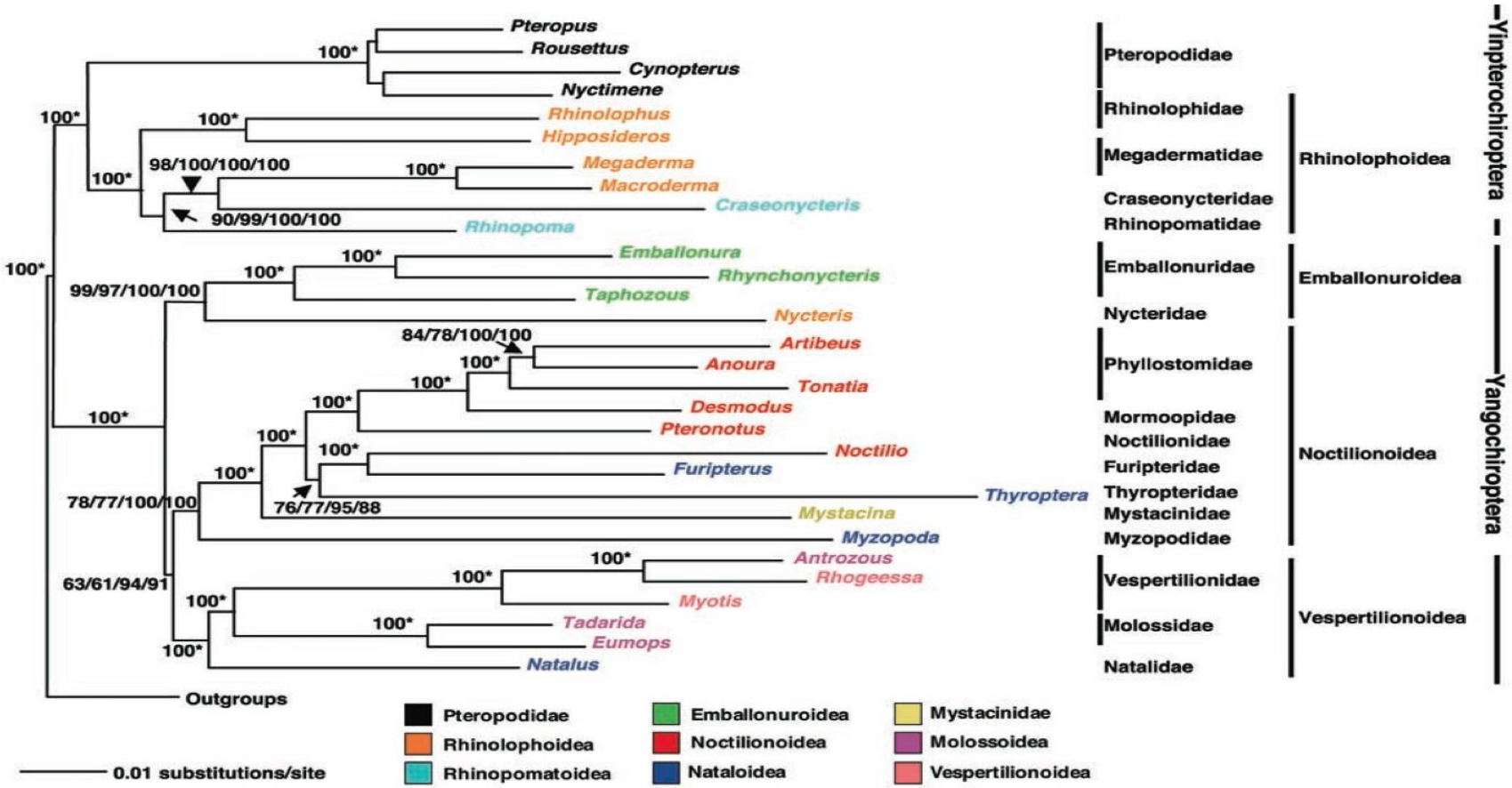


Figure 1.2: Maximum likelihood tree showing the division of the suborders Yinpterochiroptera and Yangochiroptera

From (Teeling et al., 2005). Node labels are 1. ML unconstrained bootstrap values. 2. ML constrained bootstrap values. 3. Bayesian single model posterior probabilities, (%). 4. Bayesian partitioned model posterior probabilities, (%). Where node is labelled 100*, clades received 100% bootstrap support and posterior probabilities of 1.000. Genera coloured by super-familial group.

1.6.3. Great British bat taxonomy

There are 17 species of bats in Great Britain. Two of these species, *Rhinolophus hipposideros* and *R. ferrumequinum*, are classed in the Yinpterochiroptera, and the remaining 15 species in Yangochiroptera (all family Vespertilionidae).

The increasing uptake of DNA barcoding as a tool for classifying species has led to an increase of the number of identified bat species (Clare et al., 2007). It has been particularly valuable in distinguishing anatomically cryptic species. As a result of molecular evidence it is now widely accepted that that morphologically similar pipistrelle bats: *P. pipistrellus* (Schreber, 1774) and *P. pygmaeus* (Leach, 1825) are distinct species (Barratt et al., 1997a, Hulva et al., 2004). Various studies argued that these were reproductively isolated species, using a number of molecular markers to support this (Benda et al., 2003, Mayer and von Helversen, 2001), before being accepted by the International Commission on Zoological Nomenclature in 2003 (I.C.Z.N., 2003). Additionally, a small Myotis bat, *M. alcathoe*, was discovered in 2001 in Greece (Von Helversen et al., 2001), and molecular evidence was used to confirm its presence in Great Britain in 2010 (Jan et al., 2010).

M. myotis was declared extinct in Great Britain in 1990 (Hutson et al., 2008). Individuals have been occasionally found, but as there is no evidence of breeding, *M. myotis* is considered a vagrant in Great Britain (Concannon et al., 2005, Stebbings, 1986). Other recorded vagrants include *Eptesicus nilssonii, Hypsugo savii, Myotis dasycneme, Myotis emarginatus, Pipistrellus kuhlii,* and *Vespertilio murinus* (Bat Conservation Trust, 2010). The breakdown of the species found in the UK and Great Britain are found in table 1.3. This work focuses on samples from Great Britain (England, Wales and Scotland), excluding *M. myotis*. Northern Ireland is not included as it is beyond the scope of this thesis.

Table 1.3: Breakdown of the species of bats found in England, Northern Ireland, Scotland, and Wales

Data taken from the National Bat Monitoring Programme 2014 (Bat Conservation Trust, 2015, Barlow et al., 2015), and the IUCN red list (I.U.C.N., 2013). Trends are for the UK. NT- near threatened, LC- least concern, DD- data deficient.

Species	IUCN categorisation	IUCN Population Trend	United Kingdom			
			Great Britain			
			England	Scotland	Wales	Northern Ireland
B. barbastellus	NT	Decreasing	✓		✓	
E. serotinus	LC	Unknown	✓		✓	
M. alcathoe	DD	Unknown	✓			
M. bechsteinii	NT	Decreasing	✓		✓	
M. brandtii	LC	Stable	✓		✓	
M. daubentonii	LC	Increasing	✓	✓	✓	✓
M. mystacinus	LC	Unknown	✓	✓	✓	✓
M. nattereri	LC	Stable	✓	✓	✓	✓
N. leisleri	LC	Unknown	✓	✓	✓	✓
N. noctula	LC	Unknown	✓	✓	✓	
P. nathusii	LC	Unknown	✓	✓	✓	✓
P. pipistrellus	LC	Stable	✓	✓	✓	✓
P. pygmaeus	LC	Unknown	✓	✓	✓	✓
P. auritus	LC	Stable	✓	✓	✓	✓
P. austriacus	LC	Unknown	✓		✓	
R. ferrumequinum	LC	Decreasing	✓		✓	
R. hipposideros	LC	Decreasing	✓		✓	✓

1.6.4. Methods of identifying bats

The identification of bats is crucial to their conservation. There are a number of methods available for identifying bats, each with advantages and disadvantages.

1.6.4.1. Direct methods of studying bats

The most conventional method of identifying bat species involves capture, and measurement of morphological features. These morphological features include wing measurements (length and area), tail area, forearm length, ear length, tragus (width and length), and colour (Schmieder et al., 2015, Stebbings, 1986). However, morphological

characteristics are strongly correlated with flight, and these types of measurements can confound phylogenetic relationships (Jones et al., 2002). Additionally, morphometrics cannot reliably distinguish between cryptic species such as between *P. pipistrellus,* and *P. pygmaeus*, and between the small Myotis bats *M. mystacinus, M. brandtii,* and *M. alcathoe.*

Measurement of morphological features depends upon bat capture, which requires licenced and experienced personnel. The type of licence required depends on the method used for trapping (Conservation of Habitats and Species Regulations, 2010). The benefits of capture studies are that the data are up-to-date (i.e. that there are bats present), they can provide some information on abundance, and that there is a (relatively) low cost of equipment in comparison to some other methods. The downsides of capture studies are that they rely on the experience of the personnel, the study relies on being able to catch the bats; this can be a slow process which can be adversely impacted by poor weather. Cryptic species can be difficult to differentiate, as is seen in the case of identifying *M. alcathoe*. Additionally, there may be morphological variation within a species, particularly when studying mixed sex or age populations. When handling bats there is a risk of zoonotic infection, therefore it is imperative that all bat surveyors have up-to-date rabies vaccinations.

The second invasive method of identifying species uses tissue samples, such as fur, blood, or wing biopsy. These samples are then used for genotyping by sequencing, often referred to as barcoding, which is the method of using a short genetic marker to assign an organism to a particular taxonomy, often to species level. Dependant on the barcoding region used, genotyping can distinguish between cryptic species. The disadvantages of direct genotyping are that it requires bat capture with the additional potential to affect the fitness of individuals.

1.6.4.2. Indirect methods

Acoustic monitoring of bats for taxonomic identification is, compared to the use of morphological identification, a relatively new area of study (Walters et al., 2013). Echolocation call frequencies and patterns are diverse (table 1.2), and show some connections to phylogenetic structure, although there is some overlap in frequencies used by the different species (Jones and Teeling, 2006).

The advantages of using echolocation call frequencies to classify bats are numerous. The data are up-to-date, and may provide some information on abundance. As these approaches do not require capture of bats, or other direct contact, no licenses are required. On the other hand, acoustic monitoring requires experience of the surveyor. Cryptic calls, variation within a species (males vs. females, juveniles vs. adults), and the presence of multiple species can confound species identifications. Some bats can be easily identified by audio calls, such as the rhinolophidae, whereas others, such as *Myotis* spp., are far more difficult to identify (Obrist et al., 2004). Audio detection can create vast volumes of audio file data, particularly where static, constant recording detectors are used. This can be automated, however this requires sophisticated computational approaches, which are only as good as the training set of calls, and which may not be able to identify novel species (Gaston and O'Neill, 2004).

Evidence such as discarded prey remains (rejectamenta) below perches and within droppings is sometimes used. Rhinolophid bats typically bring prey back to the roost to feed, rather than feeding on the wing. *R. ferrumequinum* dropping piles often contain beetle (Coleoptera) wing cases, whereas *R. hipposideros* dropping piles typically include the wings and legs of crane flies (Tipulidae) (Stebbings, 1986). There are only appropriate diagnostics available for some species, so this is not extensively used as diagnostic; typically, it is used in support of other methods.

1.6.4.2.1. Genotyping using guano

Guano is a readily available source of bat DNA that is typically not invasive, and can be used for genotyping using sequencing. Whist initial start-up costs of setting up a molecular biology laboratory are high, the high throughput potential of PCR and sequencing, means that costs per sample can be very low (<$5 per sample) (Hebert and Gregory, 2005), particularly in comparison to the cost of an ecologist's survey which may several ecologists working over a number of nights.

1.6.4.3. Assessing the use of guano morphology to identify bat species

Guano morphology is commonly used as a species diagnostic. This takes into account measurements such as particle size, shape, diameter, length, and colour (Stebbings, 1986). This can be done at any time, even in winter, as there are no seasonal limitations. No (or limited) is equipment needed, it does not require a license, unless the bats are being disturbed with torches, in which case a class one license is required (Conservation of Habitats and Species Regulations, 2010). However, there is a great deal of variation within a species, often as a result of the life stage of the bats: juveniles will have very different guano morphology to adults. Additionally, there is often large overlap between species (as discussed in chapter 6). The standard reference measurements are based on samples which were measured when fresh (Bat Conservation Trust, 2014 , Stebbings, 1986), and so work better with fresh samples: dry samples often shrink and change colour. Finally, the data are not current; a positive species identification does not necessarily mean that the individual is still present. The efficacy of this method, in addition to the confounding influence of diet on guano morphology is discussed in chapter six.

1.7. Scope and aims of thesis

Reconstruct dietary niche breadth of each Great British bat species- all bat species within the Great Britain are insectivorous, although some also eat arachnids (Clare et al., 2011). By forming a picture of how insect fauna is partitioned between bat species, it will be possible to ascertain to

what extent there is dietary niche overlap. By studying all of the bat species present in Great Britain under one methodological “umbrella”, ascertainment biases will be reduced. By developing new primers to expand the range of target species amplified during metabarcoding from previous studies, a better picture of the dietary taxa will be formed. Undertaking shotgun metagenomics, will allow metadata on the samples to be collected, as well as potentially identifying taxa which do not amplify using metabarcoding. This will test how effective the primers used for the metabarcoding are at amplifying the target species, and allow the two methods to be compared and contrasted. The overarching aims of this thesis are:

1. Establish generalists and specialists- In general, bats with a higher extinction risk have a narrower dietary breadth (Boyles and Storm, 2007, Jones et al., 2003). By identifying the prey species eaten by each species it will be clearer which species of bats are generalists and which are specialists. This can inform predictions about which species are most at risk of extinction, particularly in the face of changing arthropod distributions in a changing climate.

2. Investigate temporal and spatial dietary variation- By selecting samples from across the breadth of Great Britain we hope to be able to study seasonal and spatial variation within species and between species. This will give indications of the potential implications of environmental changes such as climate change.

3. Evaluate complimentary methodological approaches- By using both metabarcoding and a shotgun metagenomic approaches, the efficacy of each method will be evaluated.

1.7.1. Research hypotheses

1. Sympatric species will have highly diverse dietary preferences. This will mean that competition between these species is avoided.

2. Dietary diversity and extinction risk are correlated. Species with narrow dietary diversity are hypothesised to be those species most in decline (Boyles and Storm, 2007, Safi and Kerth, 2004). Population declines in *B. barbastellus*, *M. bechsteinii, R. ferrumequinum,* and *R. hipposideros,* (table 1.3) may be a result of narrow dietary diversity.

3. Metabarcoding will provide a high-resolution diet for each of the bats. This approach provides more DNA sequence data specific to Arthropoda than shotgun metagenomics.

4. Shotgun metagenomic data will capture a wider range of diet species than metabarcoding and provide data on gut microbiota. Shotgun metagenomic data will have a smaller number of DNA sequences assigned to Arthropoda than in the metabarcoding method, although there will be a lower number of PCR-derived duplicate sequences.

5. Dietary niches will recapitulate phylogeny. The dietary niches will have evolved with each species.

4. Variations in diet will confound guano morphology. The proportions of hard Arthropoda prey will have direct impact on the particle size of the guano. The diet will also impact the colouring of the guano.

1.7.2 Outline of thesis

The main bulk of this project falls into four sections:

Section one: A meta-analysis of the literature that pertains to bat diets. These data will be used to inform the design of primers in the barcoding stages of the project. This is discussed in chapters one, two, five, and seven.

Section two: A shotgun metagenomic analysis of a selection of guano samples from across the range of Great British species. This method will provide information, not only about diet species, but also about the bat, viral, and bacterial DNA. This is discussed in chapters three, five, and seven.

Section three: A targeted amplicon metagenome study of the mitochondrial cytochrome c oxidase subunit 1 barcode region from the arthropod species identified in the literature review, and from shotgun metagenomic data. This is to provide a greater resolution picture of the diet species present. I discuss this in chapters four, five, and seven.

Section four: In chapter five, I will summarise the data. Chapter six will look at the appropriateness of using guano morphology as a diagnostic of species presence. In chapter seven I will place the data in an ecological context.

Chapter 2 : Great British bats form four dietary niches: a meta-analysis of bat diet data (1929-2016)

2.1. Abstract

The diversity of bat species in Great Britain is thought to be as a result of extensive resource partitioning. By forming a picture of how arthropod prey are partitioned between bat species, it becomes possible to determine the level of dietary niche overlap. The information taken from the published literature will form the foundation of further analysis in this thesis.

In this chapter, I review the methods typically used to measure bat prey species. I then compile the published data on bat diets in order to describe (i) the dietary species richness and breadth, and (ii) the dietary overlap between bat species. The data shows that molecular methods identify the highest count of prey taxa, potentially to genus level, although this technique is currently limited by the reference databases available. Lepidoptera, Diptera, and Coleoptera comprise a large proportion of the diets, and are responsible for the formation of three major dietary guilds.

Some common species, such as *P. auritus*, were found to have surprisingly narrow dietary preferences. However, other, less common species, such as *M. bechsteinii,* were found to have surprisingly broad diets, a finding that is at odds with current opinions about extinction risk and dietary preferences. This demonstrates the need for all of the species to be studied using one standardised system using molecular methods. Data from this work, alongside the metagenome data from chapter three, were used contribute to primer design for the metabarcoding in chapter four.

2.2 Introduction

In 1997, Vaughan undertook a review of all of the published diets of the 15 species of bats then known to be present in Great Britain (Vaughan,

1997). Since then, two new species of bat have been described, and have been identified to be present in Great Britain: *P. pygmaeus* and *M. alcathoe*. Furthermore, 58 additional studies have since been published. It was, therefore, deemed valuable to produce an up-to-date synthesis of the all data pertaining to the diets of Great British bats.

2.2.1. Rationale for measuring bat trophic ecology data

In general, bats with a higher extinction risk are those with greater dietary specialisation (Boyles and Storm, 2007, Jones et al., 2003). Where a bat has a narrow dietary range, it is considered more vulnerable to extinction, whereas a bat with a broad dietary breadth is considered to be more robust. Knowledge of the level of dietary specialisation of a species, coupled with intelligence of the conservation status of the prey species, is a valuable resource.

Dietary overlaps will be used to describe guilds; subgroups of the community, characterised by a distinct behavioural characteristics such as diet, habitat selection, and feeding style (Patterson et al., 2003, Schnitzler et al., 2003). Understanding dietary overlaps, and the basis of the diets, will be valuable in helping to direct conservation efforts, and in explaining inter-species interactions.

Prey availability may, in some circumstances, be a more important determiner of species ability to co-inhabit in an area than dietary differences (Griffiths, 1975). If a prey species is rare, it may become a limiting factor, thus introducing competition between those predating upon it. If a prey species is common, competitive exclusion can cause one bat species to be excluded (Hardin, 1960).

2.2.2 Methodologies employed in measuring diet

There are a range of different methods employed to characterise bat diets. The first relies on direct observation of bat feeding habits (Poulton, 1929). This method provides no information as to the proportion of prey types consumed: it gives only presence data. It is also requires a great

deal of expertise to be able to accurately identify prey. Indeed, identifications made in this manner are rarely to a lower taxonomic level than order (Poulton, 1929). Second, is the analysis of prey remains found beneath known bat perches (the rejectamenta) (Mikula and Čmoková, 2012). This allows a greater resolution of prey identification, often to species level (Robinson, 1990, Rostovskaya et al., 2000). However, this method is limited to studying rhinolophid bats and others who will perch to consume prey: it is not useful for species such as pipistrelles, which are known to consume their prey 'on-the-wing' (Kalko and Schnitzler, 1993). It is also difficult to determine proportions of prey consumed by this method, and smaller or soft prey are more likely to be eaten whole, leaving no rejectamenta. The third method is the analysis of stomach contents (Sologor, 1980), although this method is rarely employed due to its invasive nature. The fourth method of analysing diets is stable isotope analysis. This method can be used to determine trophic niche width, however, it provides no information about the taxa consumed as insect isotopes do not vary widely between species (Bearhop et al., 2004).

The final two methods rely on guano analysis to determine the diets, and are, by far, the most regularly used methods. The first of these is the dissection of guano and the identification of prey under a microscope (Shiel et al., 1997). Morphological methods can provide information about the life stages of the insects (Krüger et al., 2013b), however, bats thoroughly masticate and digest their prey, often discarding the harder to digest fragments such as the carapace or elytra (Bohmann et al., 2011, Rabinowitz and Tuttle, 1982, Zeale et al., 2011). This increases the likelihood of misidentification, and identifications made in this manner are rarely more specific than order level (Clare et al., 2009).

The second is the molecular identification of prey DNA extracted from guano (Zeale et al., 2011). This technique is theoretically capable of identifying all prey species present, however, in practice; it is limited by the reference databases available, although this information is increasing

rapidly. Approaches for dealing with 'patchy' databases are discussed in chapter four.

The goals of this chapter are to (1) review the methods used to describe bat diets; (2) describe inter-specific dietary diversity and niche overlaps for each Great British bat as described in the literature; (3) to provide a list of required taxa used when designing primers for the metabarcoding in chapter three.

2.3. Methods

2.3.1. Selection of studies

Google Scholar and the Web of Knowledge databases were mined for published bat diets up to January 2016. The search terms combined iteratively "Diet", "Food", and "Prey", with "Bat", "Chiroptera", and the names of each of the species names, using the Boolean "AND".

In some cases, it was appropriate to exclude certain studies from the statistical analyses (see table 2.1). Typically, this was due to the reporting of diets with only presence/absence data. As it was not possible to convert these to numerical data without risking the introduction of bias, these were excluded. For the purpose of this study, the data are expressed in number of diets rather than number of studies, as each study may present more than one diet. They may present the diets of different species, they may present the diets of one species measured in different ways, or they may present diets of one species from different sampling locations or seasons.

Table 2.1: Reasons for excluding dietary study from analysis

Reason for exclusion	Notes
Wrong subject species (not bat)	
Wrong subject species (bat not present in Great Britain)	
Does not detail diet break down	E.g. Stable isotope analysis
Controlled experiments (i.e. captive fed bats)	
Not using primary data	Primary source used instead
Non-numerical (i.e. presence/absence) data	

Studies using stable isotope analysis were excluded as they do not identify the prey taxa consumed, only identify the trophic breadth. Study details are presented in appendix E.2.1.

2.3.2. Analysis of data

Once a count of the taxa identified by each study had been recorded for figure 2.1, the results of each publication were collapsed to order level to allow comparison of the studies. Figures created using R version 3.2.2. (R Developement Core Team, 2013), with clustering for figure 2.2 undertaken using hclust (R package) complete-linkage clustering.

2.3.2.1. Calculating dietary diversity and niche breadth

The Shannon-Weaver diversity function, eqn 1 (Shannon and Weaver, 1948), is a measure of the alpha-diversity of a community.

$$H' = -\sum_{i=1}^{S} p_i \ln p_i$$

(Eqn 1, Shannon-Weaver Diversity Index),

Where *s* is the richness of the community (i.e. total number of prey orders) and p_i is the proportion of *s* made up by order *i.* Where a bat species feeds on one prey type alone, i.e. just feeding on Lepidoptera, the Shannon-Weaver diversity index would be 0.

To identify generalist and specialist species, one must calculate the trophic niche breadth of the bat; which is calculated using Levin's index (Levins, 1968):

$$B = \frac{1}{\Sigma p_i{}^2}$$

(Eqn 2, Levin's Index),

This is standardised on a scale of 0 to 1 where n is the number of orders:

$$B_A = \frac{B - 1}{n - 1}$$

(Eqn 3, Levin's Standardised Index),

As with the Shannon-Weaver index, where a bat feeds on one prey type alone, the Levin's measure would be 0.

2.3.2.2. Determining trophic niche overlap

The extent to which two species overlap in their reliance on their prey sources can be measured using Pianka's overlap index (Pianka, 1973). This identifies potential competition for prey, and is calculated as follows:

$$O_{jk} = \frac{\sum_i^n p_{ij} p_{ik}}{\sqrt{\Sigma p_{ij}^2 \Sigma p_{ik}^2}}$$

(Eqn 4, Pianka's Overlap Index),

Where *j* is the first bat species diet and *k* is the second. *Pij* and *Pik* are the proportions of bat diets *j* and *k* with prey orders *i* present. Values are on a scale of 0 to 1, and a value of >0.6 is considered to be "biologically significant" (Bethea et al., 2006, Pianka, 1973, Pianka, 1981).

2.4. Results

In total, we identified 80 published studies, which passed the selection criteria (table 2.1). These contributed 216 different diets (with a breakdown shown in table 2.2). Of these diets 57 were collected in Great Britain, and the rest (159 diets) were gathered elsewhere (appendix E.2.1).

Table 2.2: The species studied and the number of diets reported

Species	Number of Diets
B. barbastellus	11
E. serotinus	26
M. alcathoe	4

M. bechsteinii	5
M. brandtii	3
M. daubentonii	13
M. mystacinus	5
M. nattereri	13
N. leisleri	19
N. noctula	7
P. auritus	26
P. austriacus	11
P. nathusii	4
P. pipistrellus	17
P. pygmaeus	5
R. ferrumequinum	23
R. hipposideros	24

Analysis is based on all of the data available, as the British subset does not cover the breadth of the Great British species. Analysis of the Great British subset can be found in appendix B.2.1. The trends seen are the same for both the subset and the full dataset.

2.4.1. Comparing the methods used to measure diet

Figure 2.1 shows the number of prey species identified by each study over time (A) and also the variability in the number of prey identified by each of the methods (B). The changes in dietary resolution may be due to a number of reasons, for example, improvements in techniques, reference materials, or equipment. The analysis of stomach contents would be expected to be as successful, if not more so, than guano dissection due to the incomplete digestion of the sample (compared to post excretion). However, this does not appear to be the case here, possibly due to the limited number of studies analysing stomach contents.

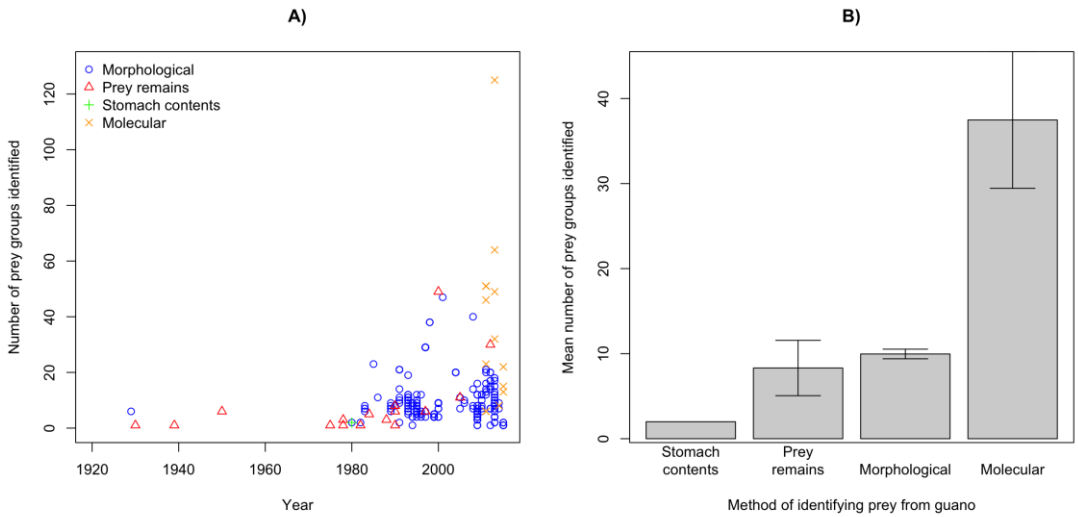


Figure 2.1: Comparison of the methods of identifying prey groups from guano
(A) The number of prey groups identified by each of the reference studies and when they were published (B) and the average number of prey groups found by each method. Code for creating figure in appendix D.2.1.

2.4.2. The diets of each bat species

There is wide variation in the primary prey consumed by the different species of bats (figure 2.2) The most important prey groups for bat species in Great Britain are Diptera, Lepidoptera, and Coleoptera. There are often clear similarities of the diets of species from the same genera shown by the clustering. This can be seen in the case of the Plecotus and Myotis species. However, the case of the rhinolophid bats, this is not seen. The *R. hipposideros* feed primarily on Diptera, whereas *R. ferrumequinum* chiefly rely on Lepidopteran species.

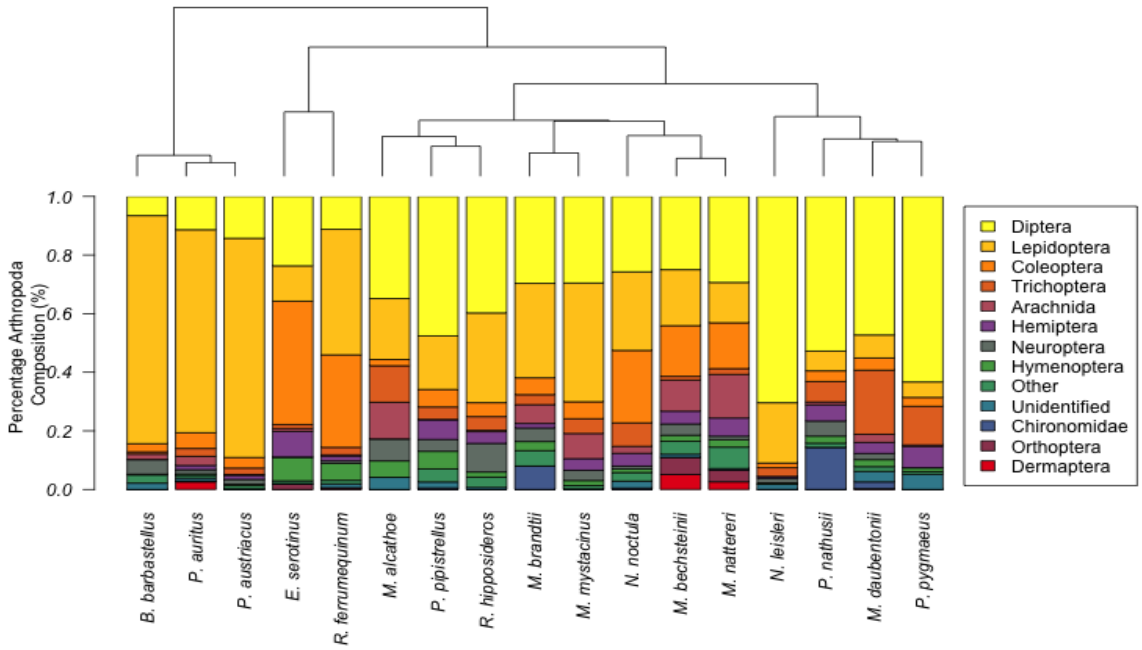


Figure 2.2: The diet of each bat species from the literature

The average diets of each of the species using all of the methods. Prey taxa have been grouped by order (where possible), class, phylum, or kingdom (where necessary). Diets sorted using complete-linkage clustering. Code adapted from appendix D.2.2.

2.4.3. Dietary diversity and niche breadth

There is wide variety in the diversity and niche breadth of the Great British bat species (figure 2.3). *B. barbastellus,* with a Shannon-Weaver diversity index of 0.935, has the smallest diversity, whereas *M. nattereri* with an index of 2.162, has the greatest diversity. The species with the smallest niche is the *P. austriacus,* with a Levin's measure of 0.0552, whereas the species with the broadest niche is, *M. alcathoe* with a measure of 0.547. The Shannon-Weaver index is sensitive to rare prey species within the diet, whereas Levin's measure gives greater weight to the dominant prey species.

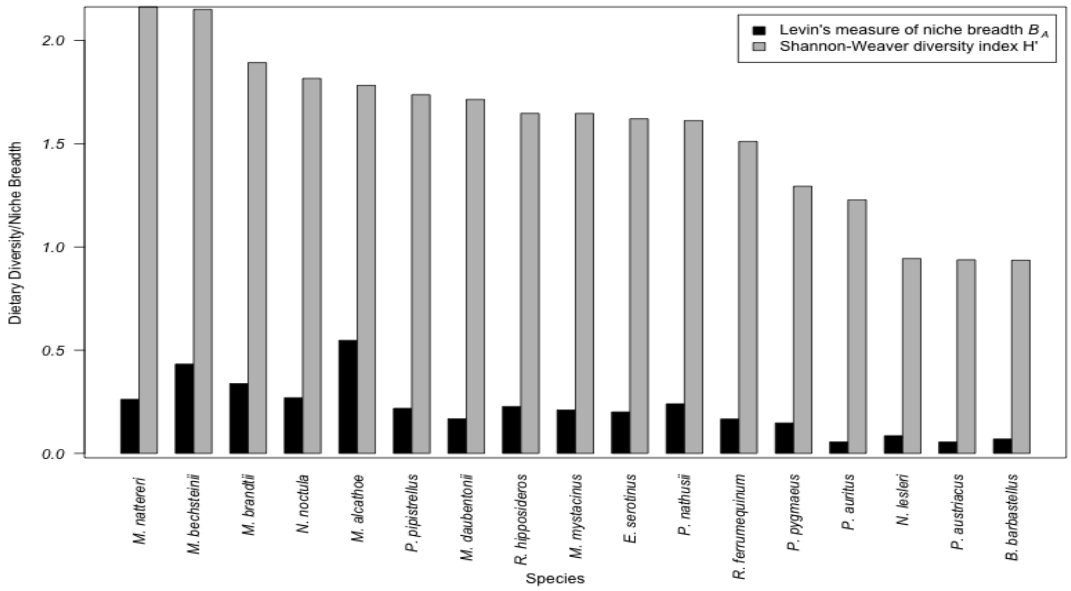


Figure 2.3: Dietary diversity and niche breadth of each species

The dietary diversity calculated using Shannon-Weaver diversity index (H) (grey) and niche breadth calculated using Levins standardised index (B_A) (black). Code adapted from appendix D.2.3.

In Vaughan's study, *R. hipposideros*, was identified as having the smallest niche breadths, whereas here, it has fallen closer to the centre of the range of diets (Vaughan, 1997).

Of the 17 species of bat in Great Britain, there are three major dietary clusters (figures 2.2 and 2.4). Guild 1, consisting of the *Plecotus* species and *B. barbastellus* have highly similar diets (Pianka's indexes of >0.99) feeding primarily (>69%) on Lepidopteran prey.

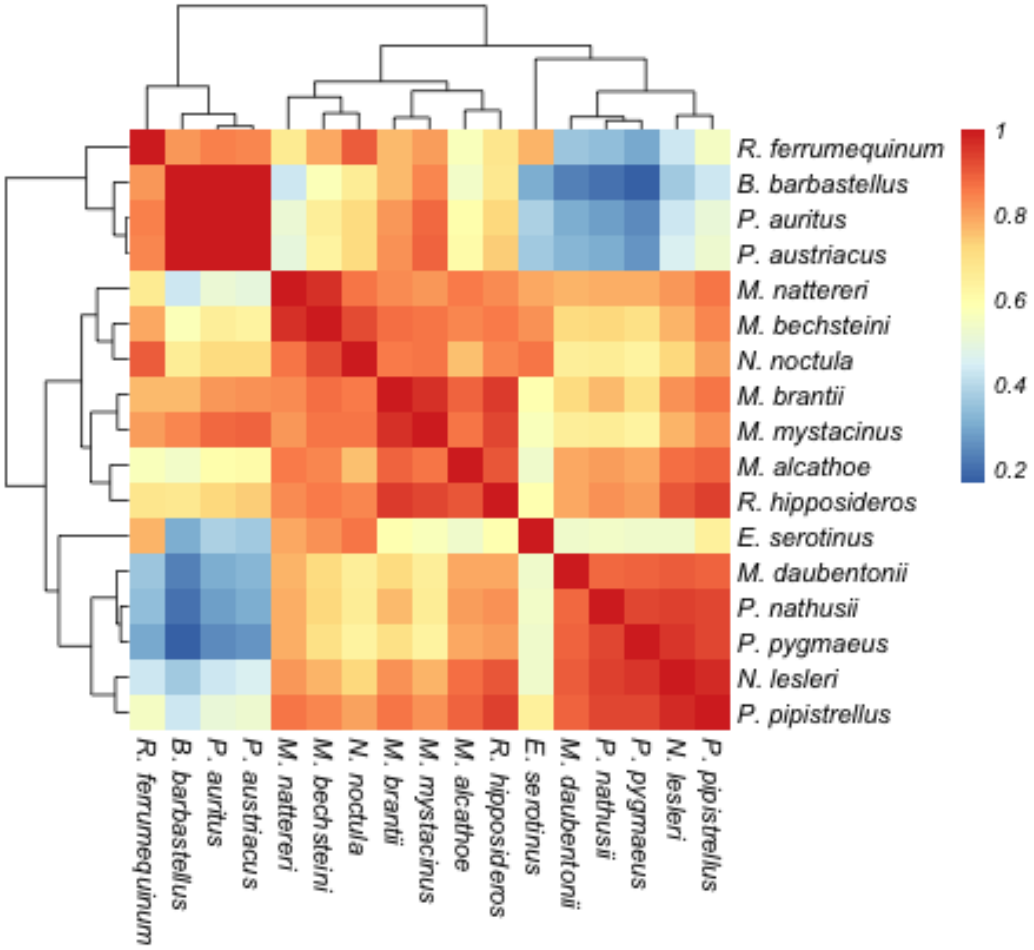


Figure 2.4: The dietary overlap of each species

Calculated using Pianka's index of niche overlap. A value of 1 (red) suggests that the diets are identical, whilst a value of 0 (blue) indicates that there is no overlap. Species have been hierarchically clustered by niche overlap similarity. Code adapted from appendix D.2.4.

The pipistrelle species complex, along with *M. daubentonii,* and *N. leiseri*, feed heavily (>47%) on Diptera. *R. ferrumequinum* and *E. serotinus* cluster poorly with the other bats, due to their high proportion of Coleoptera in their diets. The final guild feed generally on Lepidoptera, Diptera and a number of other orders. The remaining Myotis species fall in this group, along with *N. noctula*, and *R. hipposideros. R. ferrumequinum* and *E. serotinus* do not cluster into guilds.

2.5. Discussion

2.5.1. Methods used to measure bat diets

The sensitivity of molecular methods compared to other methods (figure 2.1) strongly supports the use of them for measuring bat diets. Furthermore, as databases of sequence data are developed and expanded, molecular methods would be expected to become increasingly sensitive.

2.5.2. Guild structure and niche partitioning

Where diets overlap, this suggests that there is potential for dietary competition for the species for prey, where the ranges also overlap. However, as there are around 3,000 lepidopteran species in Great Britain (Bradley and Bradley, 2000), it is necessary to have greater information about individual species consumed before one could attempt to quantify this potential competition.

The first guild, comprised of *B. barbastellus, P. austriacus* and *P. auritus*, feed primarily on Lepidoptera, and have very similar, quiet, echolocation calls (table 1.2) (Stebbings, 1986). However, their emergence times vary significantly (table 1.1), with *B. barbastellus* emerging at ~19.5 minutes after sunset (Russo et al., 2007), *P. austriacus* emerging at ~30 mins after sunset (Middleton et al., 2014), and *P. auritus* emerging at ~54 mins after sunset (Jones and Rydell, 1994). As a result, direct competition between the bats will be avoided. At these different times, it is likely that the prey species available will be different, which is further reflected in their different feeding styles (table 1.1): *B. barbastellus* is an aerial hawker (Holderied and Von Helversen, 2003), whereas *P. auritus* feeds through gleaning (Coles et al., 1989). They appear to have a narrow dietary diversity (figure 2.3), which may be an artefact of the fact that many studies do not distinguish between Lepidoptera species.

The second guild, which is comprised of the three pipistrelle species, *M. daubentonii,* and *N. leisleri,* which feeds predominantly on Diptera, has a wider range of echolocation call types than seen in guild one. Guild two

does have a broad range of emergence times. *P. nathusii* emerges between 11 and 55 minutes (Gelhaus and Zahn, 2010), *N. leisleri* at ~18 minutes (Jones and Rydell, 1994), *P. pygmaeus* at ~24 minutes (Davidson‐Watts and Jones, 2006), *P. pipistrellus* at 32 minutes (Jones and Rydell, 1994), and *M. daubentonii* at 84 minutes (Jones and Rydell, 1994). Again, this is likely a mechanism by which competition is avoided.

The third guild includes *M. alcathoe, M. bechsteinii, M. brandtii, M. mystacinus, M. nattereri, N. noctula,* and *R. hipposideros.* This guild has the highest average dietary diversity, with all species relying on a range of Lepidoptera, Diptera, and other orders.

M. alcathoe and *R. ferrumequinum* do not cluster well into guilds as defined by Pianka's indexes, however they both have high proportions of Coleoptera in their diets, which is not as heavily represented in the other bat diets. However, as the rest of their dietary preferences are dissimilar, they do not cluster with each other.

2.5.3. Niche partitioning in cryptic sympatric species P. pipistrellus and P. pygmaeus

These data do not distinguish between the two species until they were formally identified as separate species in 1999 (Jones and Barratt, 1999), with diets before this time typically being assigned to *P. pipistrellus.* In order to study the true differences between the species, it was necessary to look at the data produced since 1999 in isolation.

Using the all of the data the overlap (Pianka's overlap index) between the species was 0.926, however, when only using the data collected after 1999, the overlap was 0.806. This suggests that the diets of *P. pipistrellus* and *P. pygmaeus* may be less similar than previously thought. This may be a result of misassignment of diets to *P. pipistrellus* before the description of *P. pygmaeus* as a separate species. This further highlights the need for a rigorous comparative study of bat diets.

2.5.4. Mechanisms of trophic partitioning

There are numerous mechanisms by which trophic partitioning may occur. Species may be partitioned by habitat preference: where the primary foraging habitats are species specific, species are able to co-exist within the same ecosystem (Arlettaz, 1999). Additionally, one foraging habitat may be partitioned temporally: the emergence time of the different species of bats, plays an important part in the resource partitioning between the species (Adams and Thibault, 2006).

2.5.4.1. Feeding style and prey availability

The presence of non-volant (ametabolous) arthropods within the diet, such as Arachnida, Chilopoda, Entognatha, and Opiliones, is used to confirm the use of gleaning as a feeding style, as non-volant arthropods are unlikely have been captured in the air (Swift and Racey, 2002). In the studies surveyed, a number of bats fed heavily upon Arachnida (*M. alcathoe, M. bechsteinii,* and *M. nattereri* have Arachnida as >10% of their diets). This suggests that these are gleaning bats, unless arachnids have been caught whilst 'on-the-silk'. Both *M. bechsteinii* (Fenton and Bogdanowicz, 2002, Petrov, 2006, Wolz, 1993) and *M. nattereri* (Arlettaz et al., 1997, Geisler and Dietz, 1999, Jones, 1993, Shiel et al., 1991, Siemers and Schnitzler, 2000, Vaughan, 1997) have previously been reported as gleaning (table 1.1), whereas, *M. alcathoe* did not have any data on feeding styles. In contrast to previous literature, the rhinolophids (Ahmim and Moali, 2013, Fenton, 1997, Jones et al., 1995, Jones and Rayner, 1989) and *P. auritus* (Coles et al., 1989) only had a small proportion of their diets attributed to gleaned prey. McAney et al. reported the non-volant Siphonaptera in the diets of *E. serotinus* (McAney, 1991) which were likely ingested during grooming rather than gleaning (Shiel et al., 1998).

2.5.4.2. Bat morphology

The morphology of the bat species will impact greatly the bat's prey preference (Andreas et al., 2013, Freeman, 1979). Bat size, as well as durophagy, is correlated with prey hardness, with larger bats able to feed

on harder prey (Freeman, 1981, Freeman and Lemen, 2007, Ghazali and Dzeverin, 2013). The average Coleoptera is 3.2 times harder than moths of the same size (Freeman and Lemen, 2007), making Coleoptera amongst the hardest of the bat prey. The large bat species *E. serotinus, R. ferrumequinum,* and *N. noctula* (these bats have wing spans of about 330-450mm (Stebbings, 1986)), consume the highest proportion of Coleoptera: 42%, 31% and 24% of their diets respectively. This will be discussed in greater detail in chapter seven.

2.5.5. Trophic breadth and extinction risk

In terms of dietary specialisation, it would be expected that *M. bechsteinii*, being the greatest generalist, would be the most robust against variations in prey availability due to it's diverse diet and therefore be more successful than other species (Boyles and Storm, 2007). However, it's vulnerable status, and decreasing population belies this (I.U.C.N., 2013), showing that factors other than diet do affect population declines. Conversely, *P. auritus*, which is common in Europe, with a stable population, has a far lower prey species richness and dietary breadth, despite being predicted to have a broader dietary diversity (Battersby, 2005, Greenwood et al., 1996). This is further confounded by the limited number of studies carried out in Great Britain; diets of bats vary within species across different countries (Shiel et al., 1998). Measuring the diets of bats from guano collected from within Great Britain will be directly of use for conservation in Great Britain.

2.6. Conclusions

Great British bats form three major guilds according to their feeding preferences and dietary breadths. However, data from Great Britain, and data on population trends within Great Britain, are limited. There is need for a greater number of studies looking into temporal, seasonal and life history variations between the diets of one species. In chapter four the data from the studies will be used to direct the design of primers for metabarcoding. This will allow the diets of all of the Great British bat species to be studied under one methodological "umbrella". These data

will allow the address of areas of data deficiency in the literature, such as the feeding styles of *M. alcathoe* and *P. austriacus.* We will examine the extent to which 'known' feeding styles are reflected in the diets. This is crucial as currently there are contradictory data in the literature, particularly in regards to gleaning bats (section 2.5.4.1).

Chapter 3 : Shotgun metagenomic analysis of guano DNA

3.1. Abstract

This chapter is a shotgun metagenomic analysis of a selection of guano samples from across the range of Great Britain's species. Shotgun metagenomic methods provide information, not only about diet species, but also about the bat, viral, fungal, and bacterial DNA.

From the two Illumina runs of the shotgun metagenomic samples, 70.3% of the returned sequence data was assigned to bacterial sources, 11.8% of the sequences originated from Chiroptera, 5.85% from other vertebrates (possibly including un-mapped bat reads), 9.09% from fungi, 0.859% from viruses, and 1.16% from Arthropoda. 19 orders of Arthropoda were identified from the bat guano, of which Lepidoptera and Diptera were the highest represented. These contribute to the formation of two major dietary guilds.

The hibernation cycle appears to be a key driver in the bacterial and fungal flora associated with bat guano. The increased consumption of arthropods in order to gain fat reserves for winter appears to drive an increase in the proportion of chitin producing bacteria, and causes a skew in the ratio of Bacteroidetes and Firmicutes. During hibernation, the fall in body temperature causes the bats to have increased susceptibility to fungal infections. Bacterial infections seem to be limited to individual bats, and there is no indication from these data that there are wide spread infections of any bacterial species.

Data from the first Illumina metagenome run were used to design the metabarcoding primers used in chapter four.

3.2. Introduction

Morphological analysis of guano to determine dietary species provides a poor resolution in comparison to molecular methods (Hope et al., 2014). Stable isotope analysis has been shown to be useful for differentiating between trophic niches; prey from different trophic levels have differing enrichment of $\delta^{15}N$ (‰) (the ratio between ^{15}N and ^{14}N) (Bearhop et al., 2004, Hopkins and Ferguson, 2012). However, they do not provide information about prey species consumed. These methods are discussed at greater length in chapter two (section 2.2.2).

There have been a number of previously published studies using molecular methods, all of which have been metabarcoding based, using the primers ZBJ-ArtF1c and ZBJ-ArtR2c (Zeale et al., 2011) which target 167bp of the COI mitochondrial region. The first study, Zeale *et al,* uses cloning and traditional Sanger sequencing to identify the diet species present. This approach can be limiting, as the number of species identified cannot exceed the number of sequencing reactions undertaken, thus it rapidly becomes not cost-effective. As a result, next generation sequencing is increasingly used (Clare et al., 2011, Razgour et al., 2011). As discussed in chapter one, there are a number of limitations to metabarcoding, hence the need for shotgun metagenome data (Bon et al., 2012, Srivathsan et al., 2015).

3.2.1. Selection of sequencing method

There is a range of options available for DNA sequencing. Sanger sequencing remains the accepted gold standard in sequencing, in terms of accuracy (99.999%), with possible read lengths of >1000. Whilst the cost per Sanger sequencing reaction is now low ($1-2) the cost per 1000 bases sequenced is extremely high (~$500 per megabase) (Shendure and Ji, 2008, Wetterstrand, 2015). Additionally, methods of sequencing which rely on DNA amplification through PCR, such as Sanger sequencing, can be problematic when working with poor quality and potentially degraded DNA such as ancient DNA (aDNA) or faecal DNA,

as DNA fragmentation can prevent primer extension, and thus PCR amplification.

Next generation sequencing is considerably cheaper and faster than Sanger sequencing, with a greatly reduced hands-on sample preparation time (Hert et al., 2008). Additionally, the number of PCR cycles required on any target molecule is greatly reduced, which helps to avoid the introduction of PCR biases. Read-length is the key limitation of many next generation sequencers, although the read length chemistry is being increased all the time. Furthermore, initial input costs are much higher for running a next generation sequencer (Shendure and Ji, 2008). Computationally, the hugely increased volume of data produced by next generation sequencing provides a far greater challenge to the bioinformatician than Sanger sequencing data (Scholz et al., 2012).

The impact of next generation sequencing (NGS) on genetics is broad; it is now feasible to sequence whole genomes, transcriptomes, metagenomes (and many other ‘omes), as well as the sequencing of ancient DNA (aDNA)(Hofreiter et al., 2015, Knapp and Hofreiter, 2010, Mardis, 2008b, Mardis, 2008a). At present there are two major classes of next generation sequencer; those which use massively parallel sequencing methods, referred to as second-generation sequencers, and single molecule sequencing methods, referred to as third-generation sequencers.

Second generation sequencers are characterised by their sequencing of many molecules at the same time (massively parallel sequencing), typically after the enrichment of the target DNA. Second generation sequencers include Roche/454, Illumina, and Ion Torrent sequencers (Mardis, 2008b). The major advantage of these technologies is that they generate vast amounts of data, although this can cause computational and bioinformatic challenges. However, they have short read lengths and poorer accuracy than Sanger sequencing. These technologies have been widely used for studying degraded DNA.

Third generation sequencing machines are characterised by the use of single molecule sequencing, and include PacBio Single Molecule Real-time Sequencing (SMRT) (Quail et al., 2012), and the Oxford Nanopore MinION and PromethION (Mikheyev and Tin, 2014). At present these technologies have a high error rate but often very long read-length. At present these technologies are less appropriate for degraded DNA than the second-generation sequencers as they return a lower number of sequences, and the degradation of the DNA limits the read length.

The cost of running a next generation sequencer is still relatively high; depending on the type of sequencer used the cost of a single run may be anywhere between ~$500-$3000. Sample preparation and downstream data analysis adds considerably to this cost (Sboner et al., 2011). However, the cost per megabase can be as little as ~$0.01 and is falling all the time (Shendure and Ji, 2008, Wetterstrand, 2015). Whilst the cost of sequencing is greatly reduced, the costs of computation for analysis and data storage have not dropped proportionally (Sboner et al., 2011). However, the cost is still far cheaper than that of Sanger sequencing.

Due to its cost effectiveness and read length that is sufficient for the phylogenetic classification of sequences, an Illumina MiSeq was used for both the shotgun metagenomics (chapter three), and the metabarcoding studies (chapter four). This allowed for numerous samples to be multiplexed together in one machine run, further helping to avoid bias and reduce costs. Unfortunately, PCR free methods are not suitable in this instance, due to the low amounts of DNA in each sample; the Illumina TruSeq DNA PCR-free library preparation method requires 1-2 μg input DNA per sample. The average DNA yield from the guano samples used is ~513ng (±1006), some of which has been used for bat DNA barcoding, and some will be for the metabarcoding (chapter 4).

For a MiSeq V2 reagent chemistry run with a read length of 2x 250bp, the expected data yield is between 7.5 and 8.5 GB. Assuming that 7.5 GB is

returned over 108 samples (including the negative control), ~ 69.4 MB of sequence data would be returned per sample (assuming that reads are greater than 250bp). We made an assumption that 10% of the DNA is from Arthropoda, and, if so, this would give 6.94 MB of Arthropoda DNA sequences, equating to 27,760 reads (13,880 unique reads if forward and reverse reads are entirely overlapping). Previous molecular studies (explored in greater detail in chapter 2) identified an average of 37 diet taxa with (Krüger et al., 2013a) identifying the most diet taxa (125), so this amount of data should be sufficient for arthropod identification.

3.2.2. Avoiding biases

The major advantage of this study, in comparison to many other studies, is that it uses only samples collected within Great Britain, and compares all samples under identical methodologies. By using a shotgun metagenomic approach, primer biases and some PCR biases are avoided. There is no need for the *a priori* knowledge required for primer design. Additionally, due to the use of fewer rounds of PCR in the sample preparation process used for preparing a shotgun metagenome, there is less PCR stochastic bias (Polz and Cavanaugh, 1998). Shotgun metagenomics can also be used to provide metadata for the target data, such as gut microbiome data when studying diet. Direct sequencing allows identification of organisms which cannot be cultured, or cannot be distinguished from other species by using targeted sequencing such as 16S rDNA (Tringe et al., 2005). A major advantage of shotgun metagenomics is that it exploits more of the DNA laid down by the organisms than barcodes, and so has the potential to be more sensitive than metabarcoding especially with fragmented DNA (Smith et al., 2015).

There are a number of disadvantages associated with shotgun sequencing; the endogenous DNA of the target organisms can be low in comparison to other DNA sources, depending on the type of sample used. Additionally, the database coverage is typically far more restricted than that of amplicon targets. This causes a greater risk of

misassignments of sequences to over-represented (i.e. model organism) sequences in the database.

DNA capture methods were not used as they carry the same biases as the metabarcoding approaches, and are considerably more expensive, without providing significantly more information that could be used in this study.

3.2.3. Pathogens carried by Great British bats

Bats have long been identified as sources and as carriers of zoonotic infections (Han et al., 2015, Wong et al., 2007). Furthermore, there are pathogens, which, whilst they are not harmful to humans, can cause mass mortality and morbidity in bat populations, such as *Pseudogymnoascus destructans,* the causative agent of white-nose syndrome. Viruses are the best studied of bat associated pathogens, although bacteria and fungi contribute significantly to pathogenesis in bats (Mühldorfer, 2013). Fears about the spread of diseases can lead to persecution and interventions such as culling of bats, which may not be effective at controlling pathogens (Blackwood et al., 2013, Hallam and McCracken, 2011) and can decimate bat populations (Florens, 2015).

Rhinolophus associated viruses include *Hantavirus, Betacoronavirus, Flavivirus, Pestivirus, Orthohepadnavirus, Picornaviridae, Spumavirus, Rhabdoviridae*, and *Alphavirus*. Viruses that are associated with the vespertilionidae include *Mastadenovirus, Bornavirus, Hantavirus, Phlebovirus, Dicistroviridae, Flavivirus, Picobirnavirus, Nodavirus, Chiropoxvirus, Gammaretrovirus, Vesiculovirus, Picornaviridae,* and *Betacoronavirus* (Wang and Cowled, 2015). Of these, *Orthohepadnavirus* (Drexler et al., 2013), *Mastadenovirus* (Kohl et al., 2011), and *Chiropoxvirus* (Emerson et al., 2013), are dsDNA viruses. Due to the library preparation method, we will only identify dsDNA viruses (Baltimore classification group 1). In order to process other types of viruses (ssDNA, ssRNA, dsRNA) one would have to undertake additional, alternative library preparation methods, which is beyond the scope of this study.

Enteropathogenic bacteria that are associated with bats include *Campylobacter jejuni* (Hazeleger et al., 2011), *Clostridium perfringens (Hajkova and Pikula, 2007), Clostridium sordellii* (Mühldorfer et al., 2011a), *Listeria* spp. (Rozalska et al., 1998), *Salmonella* spp. (Reyes et al., 2011), *Shigella flexneri* (Rozalska et al., 1998), *Vibrio* spp. (Mühldorfer et al., 2011a), and *Yersinia* spp. (Mühldorfer, 2013, Rozalska et al., 1998). Enteric bacteria are thought to originate from the diet or foraging habitats of the bats (Mühldorfer, 2013). There are a number of bacteria that are arthropod borne pathogens which have been recorded in bats, including *Bartonella* spp. (Concannon et al., 2005, Kosoy et al., 2010, Lin et al., 2012), *Borrelia* spp. (Hanson, 1970, Petney et al., 2000), *Grahamella* spp. (Sebek, 1974), and *Neorickettsia risticii (Gibson et al., 2005). Leptospira* has also been identified in bats (Fennestad and Borg-Petersen, 1972, Matthias et al., 2005). Bacterial DNA comprises a large proportion of the DNA extracted from faeces, and so these pathogens, in particular the enteric bacteria, may be represented in the shotgun metagenomic datasets.

Pseudogymnoascus destructans (formally known as *Geomyces destructans),* is the causative agent of white-nose syndrome, and has caused catastrophic population crashes in a number of bat species in North America (Lorch et al., 2011). *P. destructans,* is known to be present in a number of countries across Europe, however it has not been associated with the mass mortality seen in North America (Puechmaille et al., 2011). *P. destructans* was introduced to America from Europe (Leopardi et al., 2015), and was found to be present in bats in Britain in 2013, although as yet there are no recorded cases of white-nose syndrome in Britain (Barlow et al., 2015a). Other fungi carried by bats include: *Histoplasma capsulatum*, the causative agent of histoplasmosis, has been identified in bat guano (Emmons et al., 1966, Miller, 1992), and a number of *Aspergillus* species, which are the causative agent of aspergillosis, can be found in bat guano (Miller, 1992, Nováková, 2009).

3.3. Materials and methods

3.3.1. Wet lab

3.3.1.1. Selecting samples

Samples were selected in order to cover all Great British species, and, where possible, selected so that they covered wide geographic range (Scotland, northern England, central England, Wales, and southern England). Samples that had been submitted at different times of the year were also chosen. This may help to give an indicator about seasonal variation; however, we do not have exact information about exactly when some of the samples were deposited. The majority of the samples have been taken from those submitted for the Ecowarwicker Ecological Forensics Bat ID service. Where there were gaps in the species coverage from the samples from the BatID scheme, the samples to be used have been supplemented by David Bullock (National Trust), and Ian Tanner (Ecoline). The locations from which the samples were taken can be seen in figure 3.1.

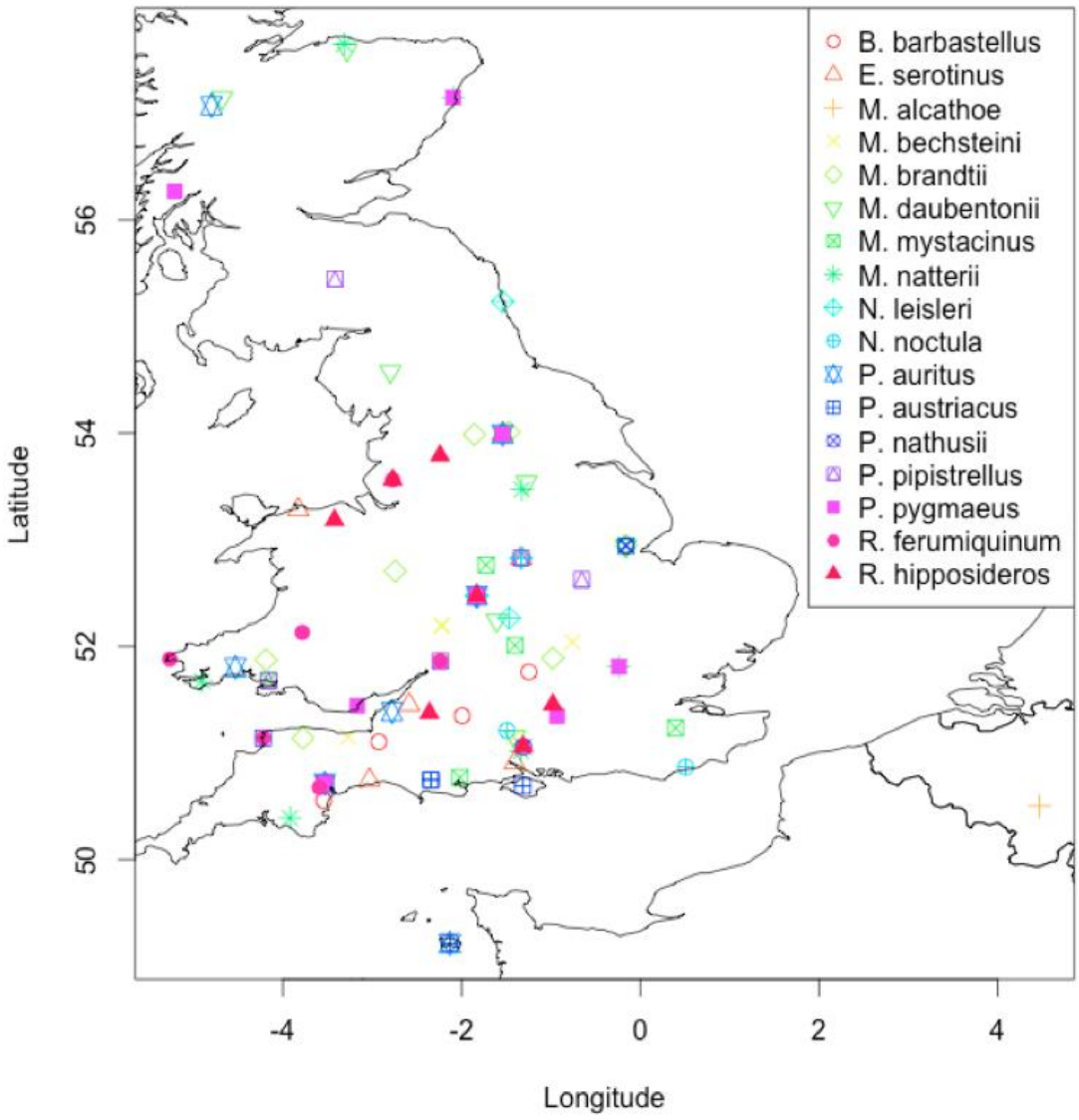


Figure 3.1: Distribution of samples

Where possible, samples from each species were selected from across the range for that species. Figure made using R packages rworldmap, rworldxtra, RColorBrewer, and plotrix (R Developement Core Team, 2013). For code, see appendix D.2.5.

Table 3.1 shows the number of libraries prepared for each species. Only one run of the Illumina MiSeq was planned for this section, however, due to a technical failure during the first round of MiSeq sequencing, only around 10% of the expected data was returned. The arthropod data from this run was used in the design of the metabarcoding primers (chapter four). As a result of the limited data from the first run along with the poor coverage of some of the species, it was decided to re-run the MiSeq,

incorporating a number of new samples, which were not available during the first run (see tables 3.4 and 3.5).

Table 3.1: Counts of libraries for each species

Each library made from one individual guano.

Species	**Number of libraries Illumina MiSeq run 1**	**Number of libraries Illumina MiSeq run 2**
Negative control (Blank)	1	1
Positive control (*P. pygmaeus*)	1	1
B. barbastellus	4	6
E. serotinus	4	6
M. alcathoe	0	1
M. bechsteinii	4	6
M. brandtii	4	7
M. daubentonii	4	7
M. mystacinus	4	7
M. nattereri	5	7
N. leisleri	4	6
N. noctula	3	6
P. auritus	7	7
P. austriacus	4	7
P. nathusii	4	6
P. pipistrellus	6	7
P. pygmaeus	5	7
R. ferrumequinum	4	6
R. hipposideros	4	7

3.3.1.2. Experimental controls

In order to validate the results, two experimental controls have been included in the first Illumina run: a positive control and a negative control. The positive control is a guano sample from a hand reared *P. pygmaeus*, which has been fed on a diet of *Tenebrio molitor* larval instars (meal-worm).

A negative control (reagents only, without any sample) was used to ensure that contamination issues were avoided (Salter et al., 2014b). Salter *et al,* showed recently, that if one simply sequenced a blank library

one would still return sequence data due to reagent contamination (Salter et al., 2014a). This would vary between kits, and even between batches of reagents. It is necessary, therefore, when preparing libraries to compare, to use the same reagents; it is inadvisable to use multiple kits, for example. It is also beneficial to run a blank library, multiplexed with the other libraries. This allows for any kit-derived sequences to be identified, and also for any contamination from the library preparation stage to be seen.

3.3.1.3. Sample preparation

3.3.1.3.1. Addressing the technical considerations involved with DNA from guano

To tackle the challenges of contamination discussed in section 1.4.4, DNA extractions, library preparations and PCR setup are undertaken in a dedicated, chambered laboratory, which is physically separated from post-PCR laboratories (Palmer et al., 2012), see appendix A.1. To avoid PCR inhibition, chloroform extractions are used to remove inhibitors (Wiedbrauk et al., 1995). Acetone washes were used to further remove inhibitors and traces of ethanol from the extraction kit (Prado et al., 1997).

3.3.1.3.2. Extraction of DNA

For each sample, one guano was placed into a microfuge tube, and crushed using a sterile plastic stick (Puechmaille et al., 2007). Only one guano per sample was used, as this typically provides a sufficient DNA yield for Illumina library preparation, and avoids the potential mixing of guanos from different individuals, in particular, from individuals from different species as is possible where guano samples were collected from mixed roosts. The crushed samples were incubated overnight at 37°C in 300 *µ*l CTAB buffer on a sample agitator at 400rpm. CTAB extractions yield higher concentrations of DNA from guano than MoBio, Epicentre, and Qiagen stool kits (Jedlicka et al., 2013). CTAB stabilises DNA, and facilitates the separation of polysaccharides when high sodium chloride concentrations are used; at high concentrations, polysaccharides are insoluble, whereas at low NaCl concentrations, DNA is insoluble (Doyle

and Doyle, 1987, Monteiro et al., 1997). Incubation at low temperatures, for extended time (compared to typical extractions at 65°C for one hour), reduces the risk of heat-induced damage to the DNA, whilst still yielding high DNA concentrations: an important consideration when working with recalcitrant DNA (Rohland and Hofreiter, 2007).

DNA was extracted using chloroform:isoamyl alcohol 24:1. After spinning, the DNA is in the aqueous phase, and proteins and polysaccharides move into the chloroform/alcohol layer, removing these inhibitors. The DNA was then purified using DNeasy columns and buffers, with an additional acetone wash and dry before elution (Köchl et al., 2005, Prado et al., 1997).

After extraction, quantification of the samples was undertaken using Qubit dsDNA HS assays to confirm extraction success. Qubit was used rather than Nanodrop, as it is more sensitive at low concentrations of DNA. Qubit dsDNA HS assays were undertaken on the extraction blanks to ensure that contamination had not occurred.

3.3.1.3.3. Robust identification of guano samples

When using guano to study bats it is crucial that the species from which the dropping used originated is robustly identified; it is possible that multiple species may be roosting together, for example, with *M. daubentonii* and the *N. noctula* have been reported to roost together (Lučan et al., 2009). As a result, a positive identification of one species that is present in a site does not mean that all of the guano on that site has come from that species. When the sample is collected from a captured bat, molecular identification is still important as some species, in particular the small Myotis species *M. mystacinus, M. brandtii* and *M. alcathoe* are difficult to distinguish morphologically (Von Helversen et al., 2001). The difficulties in identifying bat species using guano morphology are discussed in chapter 6.

After the DNA had been extracted from the guano, the species of bat was confirmed using barcoding as follows; 20 *µ*l PCRs were prepared using a mixture of all of the primers shown in table 3.2, each at 5 *µ*M. Each PCR contained 2 *µ*l 10X Platinum® *Taq* buffer, 2 *µ*l of dNTPs at 2mM, 0.8µl 50mM Mg++, 1.3 *µ*l primer mix, 0.1 *µ*l Platinum® *Taq* DNA polymerase, between 0.2-2 *µ*l of sample and 11.8-13.6 *µ*l ultrapure H_2O. Touchdown PCR was used in order to account for the differences in optimum annealing temperatures of the primers used (Don et al., 1991, Korbie and Mattick, 2008). Touchdown thermal cycling conditions were as follows: 5 mins at 95°C, followed by 10 cycles of 94°C for 30s, 57°C for 30s (decreasing by 0.1°C per cycle) and 72°C for 30s, followed by 32 cycles of 95°C for 30s, 54°C for 30s, then 72°C for 30s, followed by a final extension period of 72°C for 7 minutes.

After PCR, success was determined by running on a 2% agarose gel, stained with gel red. Clean-up was undertaken by adding 2 *µ*l of Fast-AP and 0.5 *µ*l of Exonuclease-1, then incubated at 37°C for 30mins, then 80°C for 15 mins. Forward primers (BF1-7) were used in a GATC Lightrun™ Sanger sequencing reaction. Sequences were checked from traces using CodonCode aligner, then sequences were checked against the NCBI nt database, and run through our in-house pipeline: DR WHO, which confirms species identification using maximum likelihood and bootstrap support.

Table 3.2: Primers used to confirm the identity bat species of the guano

Primers designed by Robin Allaby and Oliver Smith (unpublished).

Primer name	Orientation	Sequence
BF1	Forward	ATGACAAACAYTCGAAAATCC
BF2	Forward	ATGACAAACATTCGAAAGTMC
BF3	Forward	ATGACCAACATTCGTAAATCW
BF4	Forward	ATGACCAACATTCGAAAATCY
BF5	Forward	ATGACCMACATTCGAAAATCY
BF6	Forward	ATGACCAACATTCGAAAGTCY
BF7	Forward	ATGACCAACATTCGCAARTCY
BX1	Reverse	GTCTGMTGTRTAGTGTATGG

BX2	Reverse	RTCYGATGTGTGATGCATGG
BX3	Reverse	RTCTGATGTRTAGTGTATTGC
BX4	Reverse	RTCTGATGTRTARTGTATGGC
BX5	Reverse	RTCTGAYGTRTAGTGTATAGC
BX6	Reverse	RTCTGATRTGTAATGTATAGC
BX7	Reverse	ATCTGATGTATAATGTATWGCT
BX8	Reverse	GTCTGATGTATAGTGTATGGA
BX9	Reverse	GTCTGGTGTGTAATGTATGG
BX10	Reverse	ATCTGATGTAGTGCGCATGG

As well as confirming the species, this also ensures that the DNA extraction was successful, and that the DNA from the guano is of sufficiently high quality to amplify successfully in the bat PCR reaction.

3.3.1.4. Library preparation

3.3.1.4.1. Fragmentation of DNA

DNA was thought to be able to survive for around 10^5 years in cold and dry environments and 10^4 years in temperate environments (Poinar, 2002, Poinar and Cooper, 2000), however, more recent studies have extended this to around 750,000 years (Orlando et al., 2013), with theoretical limits being much higher (Kistler et al., 2015). However, when working with bat guano, the DNA is often highly degraded, particularly where it has been collected from humid environments, or been exposed to temperature variation (Puechmaille et al., 2007). DNA is subject to attack from endogenous nucleases, changes in pH, as well as bacterial and fungal degradation (Poinar, 2002). Previous work using the DNA extracted from our bat guano samples suggested that the bat DNA was indeed fragmented, but it is unclear how fragmented.

When inputting samples into an Illumina sequencer, they should have a median length of 500 bp (Illumina, 2013b, Illumina, 2013a, Meyer and Kircher, 2010). If samples are significantly longer than this, they can interfere with cluster formation on the flow-cell. Therefore, it is important to know the fragmentation of the DNA. Fragmentation usually uses nebulisation, sonication, or enzymatic methods (Knierim et al., 2011).

Bioanalyzer traces showed some high molecular weight DNA in most samples, with some smaller fragments. It was predicted that bacterial DNA (in particular environmental bacteria) would have high molecular weight DNA, but that Arthropoda DNA would be more fragmented due to the digestion processes.

Primers for bat and insect were designed and sourced from the literature. For *P. auritus* DNA, four forward primers and four reverse primers were designed which amplified up various length sections of the cytochrome b (CytB) gene. For Lepidoptera, three forward and three reverse primers were selected from the literature to amplify the Cytochrome c oxidase 1 (CO1) gene (table 3.3 and figure 3.2) (Hajibabaei et al., 2006, Hebert et al., 2004, Zeale et al., 2011).

Table 3.3: Primers for Chiroptera cytochrome B and Lepidoptera Cytochrome c oxidase 1 used in fragmentation study

(Hajibabaei et al., 2006, Hebert et al., 2004, Zeale et al., 2011)

Primer name	Orientation	Target	Sequence 5'-3'
P_F1a	Forward	*P. auritus*	GCCATACAYTACACATCAG
P_F1	Forward	*P. auritus*	ACCAACATTCGAAAGTCYCAC
P_F2	Forward	*P. auritus*	CAGAAACCTGAACGTRGGAG
P_F3	Forward	*P. auritus*	TCTCCGTAGATAAGCAACAC
P_R1	Reverse	*P. auritus*	GRTATCGTAGYACTCAGC
P_R2	Reverse	*P. auritus*	ATTACGGTTGCTCCTCAA
P_R3	Reverse	*P. auritus*	GTGTTGCTTTATCTACGGGAGA
P_R4	Reverse	*P. auritus*	GGGTGYAAGGGAATTATATCTAT
L_F1	Forward	Lepidoptera	AGATATTGGAACWTTATATTTTATTTTTGG
L_F2	Forward	Lepidoptera	ATTCAACCAATCATAAAGATATTGG
L_F3	Forward	Lepidoptera	GCTTTCCCACGAATAAATAATA
L_R1	Reverse	Lepidoptera	WACTAATCAATTWCCAAATCCTCC
L_R2	Reverse	Lepidoptera	ATTCAACCAATCATAAAGATATTGG
L_R3	Reverse	Lepidoptera	GCTTTCCCACGAATAAATAATA

As the presence/absence of a band was a clear enough diagnostic, primers with large differences in band size (that could be clearly

distinguished on a gel) were mixed. This resulted in five primer systems (figure 3.2).

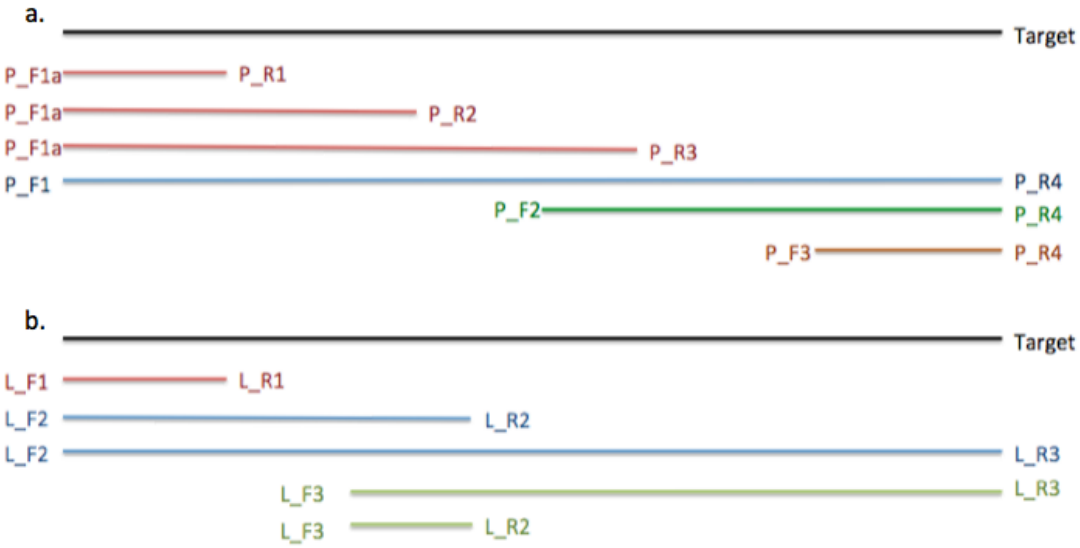


Figure 3.2: Primer system for P. auritus and Lepidoptera
(Zeale et al., 2011, Hajibabaei et al., 2006, Hebert et al., 2004) a. The primer systems amplifying *P. auritus* DNA. The total amplicon length is 648bp (P_F1a-P_R3a), and the smallest amplicon was 144bp (P_F3-P_R4) with intermediate length amplicons as follows; 520bp (P_F1a-P_R3), 416bp (P_F1a-P_R2), 319bp (P_F2-P_R4), and 224bp (P_F1a-P_R1). b. The primer systems amplifying Lepidoptera DNA. The total amplicon length is 680bp (L_F2-LR3), and the shortest was 70bp (L_F3-L_R3), with intermediate length amplicons as follows: 420bp (L_F3-L_R3), 330bp (L_F2-L_R2), and 170bp (L_F1-L_R1).

Twenty-one *P. auritus* guano samples were selected by the date that they were received; one sample every two months, from November 2009 to March 2013, to account for fluctuation of fragmentation throughout the year.

The *Plecotus* primers were successful for 86% of the samples, whereas the Lepidoptera primers were only successful in 64% of cases, suggesting that the Lepidoptera DNA is more fragmented than the bat DNA (results in appendix C.3.1). Many of the samples had bat and Lepidoptera DNA lengths of greater than the desired length for library preparation. Additionally, PCR may be amplifying the longest reads in a

distribution of fragment sizes, so may not be representative of the true fragment distribution. Furthermore, any long fragment lengths can interfere with Illumina clustering, therefore, the DNA for each of the samples used for library preparation had to be fragmented prior to undertaking library preparation. Fragmentase™ was selected to fragment the DNA. Extracted DNA was incubated at 37°C with the Fragmentase reaction mixture for 20 mins. DNA samples were then purified using 20 µl of homemade SPRI beads (see appendix A.2. and (Rohland and Reich, 2012)). After purification, they were eluted in 25 µl of EBT buffer. EBT was used for elution rather than the more usual EB, as it increases the ease of bead migration during SPRI purification due to the reduction in surface tension provided by the Tween 20 (Meyer and Kircher, 2010). The recipe for EBT can be found in the appendix A.3.

3.3.1.4.2. Preparation of adapter mix

Illumina P5 and P7 adaptors were created by incubating 200 µM of each of the appropriate adapter oligonucleotide (IS1 and IS3 for P5, IS2 and IS3 for P7 adaptor) in a thermal cycler with 1x oligonucleotide hybridisation buffer (recipe in appendix A.4). Thermal cycling conditions were 10 sec at 95°C, followed by a ramp from 95°C to 12°C at a rate of 0.1°C/sec. Both reactions were then pooled to obtain an adapter mix (100 µM each adapter). For final adaptor structure, see figure 3.3. The IS1, IS2, and IS3 adapters are modified oligonucleotides containing phosphorothioate bonds; a sulphur atom has been substituted for a non-bridging oxygen in the phosphate backbone of the oligonucleotide. The phosphorothioate bonds prevent exonucleolytic attack on the oligonucleotide (Skerra, 1992). Adapter oligonucleotides were purified using HPLC purification.

Figure 3.3: I5 and I7 adaptors after annealing

* Indicates a phosphorothioate bond.

3.3.1.4.3. Blunt-end repair

The full 25 µl of each sample (after Fragmentase™ treatment) were blunt-end repaired to fill in 5' overhangs and remove 3' overhangs, using T4 polynucleotide kinase (10 U/µl) and T4 DNA polymerase (5 U/µl) in reactions with 10X Buffer Tango, 25 mM dNTPs and 100mM ATP for a total volume of 35 µl. These were incubated in a thermal cycler for 15 mins at 25°C followed by 5 mins at 12°C. Samples were bead-purified using 35 µl of homemade SPRI beads, then eluted in 20 µl of EBT.

3.3.1.4.4. Adapter ligation

Adaptors were ligated to the samples using T4 DNA ligase (5 U/µl) in T4 Ligase buffer (10x) and Peg-4000 (50%) (see figure 3.4). These were incubated over night at 15°C, cleaned up using 40 µl of homemade SPRI beads, and eluted in 20 µl EBT buffer. Over-night incubations were used as this was found to be effective at ligating to the target DNA as described in (Tayyab, 1992), without extensive chimeric molecules being formed (which would have been seen in the PCR control library, section 3.3.1.4.6).

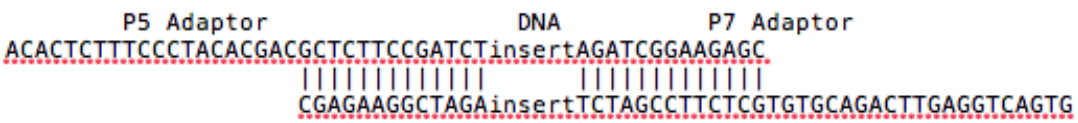


Figure 3.4: The DNA insert with adaptors ligated

3.3.1.4.5. Adapter fill-in

Adapter overhangs were filled in using *BST* polymerase (Fig. 3.5). 20 *µ*l of master mix was added to 20 *µ*l of the elutate from the adaptor ligation step to obtain reaction volumes of 40 *µ*l. These were incubated at 37°C for 20 mins, followed by 20 mins of 80°C to heat-inactivate the enzyme. These were cleaned using 40 *µ*l of homemade SPRI beads and eluted in 40 *µ*l of EBT.

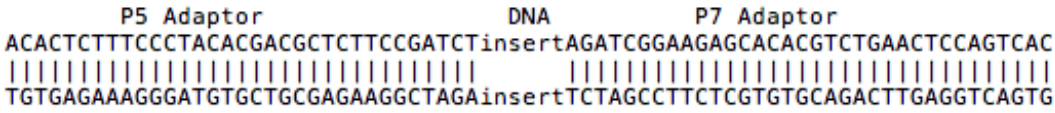


Figure 3.5: Fill-in of adaptor overhangs

3.3.1.4.6. Library characterisation

To verify the success of the library preparation protocol, electrophoresis was performed on a positive control PCR product that had undergone library preparation, and one without. The control library had shifted up by ~66bp as expected, without the formation of chimeric molecules.

3.3.1.4.7. Indexing PCR (round one)

By using uniquely indexed libraries, the libraries can later be pooled and run on one Illumina MiSeq run, and later demultiplexed, to allow the data from each library to be assigned to the appropriate sample, and analysed separately. To avoid a downstream failure of Illumina's image analysis software, subsets of indexes were selected in order to prevent unbalanced usage of the four nucleotides in any one position. In addition, indexes were designed so that they had at least 3 nucleotides different to any other index (Hamming distance of 3) (Steane, 1996), in order to prevent miss-assignment of reads due to sequencing errors. The indexing PCR process can be seen in figure 3.6. Each index (one in each adaptor) is 7 base pairs long, with an odd number, which helps to avoid palindromes. Each has an even GC content and avoids base repeats.

Indexing primers were purified using RPC purification. Details of indexing primers can be found in table 3.4 and table 3.5.

Table 3.4: Indexes used in the indexing PCR and for de-multiplexing libraries in the first Illumina MiSeq run

Indexes		D701	D702	D703	D704	D705	D706	D707	D708	D709
		ATTACTCG	TCCGGAGA	CGCTCATT	GAGATTCC	ATTCAGAA	GAATTCGT	CTGAAGCT	TAATGCGC	CGGCTATG
D501	TATAGCCT	*P. auritus*	*P. pipistrellus*	*M. mystacinus*	*M. nattereri*	*M. daubentonii*	*P. austriacus*	*P. nathusii*	*M. bechsteinii*	*N. leisleri*
D502	ATAGAGGC	*P. auritus*	*P. pipistrellus*	*M. mystacinus*	*M. nattereri*	*M. daubentonii*	*P. austriacus*	*P. nathusii*	*M. bechsteinii*	*N. noctula*
D503	CCTATCCT	*P. auritus*	*P. pipistrellus*	*M. mystacinus*	*M. nattereri*	*M. daubentonii*	*P. austriacus*	*P. nathusii*	*N. leisleri*	*M. bechsteinii*
D504	GGCTCTGA	*P. auritus*	*E. serotinus*	*P. pygmaeus*	*M. nattereri*	*M. daubentonii*	*R. ferrumequinum*	*B. barbastellus*	*N. leisleri*	Blank
D505	AGGCGAAG	*P. auritus*	*E. serotinus*	*P. pygmaeus*	*R. hipposideros*	*M. brandtii*	*R. ferrumequinum*	*B. barbastellus*	*N. noctula*	*P. auritus*
D506	TAATCTTA	*P. pipistrellus*	*E. serotinus*	*P. pygmaeus*	*R. hipposideros*	*M. brandtii*	*R. ferrumequinum*	*B. barbastellus*	*N. noctula*	*P. austriacus*
D507	CAGGACGT	*P. pipistrellus*	*E. serotinus*	*P. pygmaeus*	*R. hipposideros*	*M. brandtii*	*R. ferrumequinum*	*B. barbastellus*	Positive control	*P. auritus*
D508	TGACTGAC	*P. pipistrellus*	*M. mystacinus*	*M. nattereri*	*R. hipposideros*	*M. brandtii*	*P. nathusii*	*M. bechsteinii*	*N. leisleri*	

Table 3.5: Indexes used in the indexing PCR and for de-multiplexing libraries in the second Illumina MiSeq run

Indexes		D701	D702	D703	D704	D705	D706	D707	D708	D709
		ATTACTCG	TCCGGAGA	CGCTCATT	GAGATTCC	ATTCAGAA	GAATTCGT	CTGAAGCT	TAATGCGC	CGGCTATG
D501	TATAGCCT	*B. barbastellus*	*P. pipistrellus*	*B. barbastellus*	*M. nattereri*	*B. barbastellus*	*B. barbastellus*	*E. serotinus*	*M. bechsteinii*	*R. ferrumequinum*
D502	ATAGAGGC	*P. auritus*	*P. pipistrellus*	*E. serotinus*	*M. nattereri*	*M. daubentonii*	*P. austriacus*	*E. serotinus*	*E. serotinus*	*R. ferrumequinum*
D503	CCTATCCT	*E. serotinus*	*E. serotinus*	*M. alcathoe*	*M. bechsteinii*	*M. bechsteinii*	*M. bechsteinii*	*M. bechsteinii*	*N. leisleri*	*M. bechsteinii*
D504	GGCTCTGA	*M. brandtii*	*M. brandtii*	*M. brandtii*	*M. brandtii*	*M. daubentonii*	*M. nattereri*	*M. daubentonii*	*M. daubentonii*	*R. hipposideros*
D505	AGGCGAAG	*M. daubentonii*	*M. daubentonii*	*M. mystacinus*	*M. mystacinus*	*M. brandtii*	*M. mystacinus*	*M. mystacinus*	*M. mystacinus*	*R. hipposideros*
D506	TAATCTTA	*M. mystacinus*	*M. mystacinus*	*M. nattereri*	*M. nattereri*	*M. brandtii*	*M. daubentonii*	*B. barbastellus*	*M. nattereri*	*P. auritus*
D507	CAGGACGT	*P. pipistrellus*	*M. nattereri*	*N. leisleri*	*N. leisleri*	*N. leisleri*	*N. leisleri*	*B. barbastellus*	*N. leisleri*	*R. hipposideros*
D508	TGACTGAC	*N. noctula*	*N. noctula*	*N. noctula*	*N. noctula*	*M. brandtii*	*N. noctula*	*N. noctula*	*P. auritus*	*R. hipposideros*
D509	GTCACATG	*P. auritus*	*P. auritus*	*P. auritus*	*P. auritus*	*P. austriacus*	*P. austriacus*	*P. austriacus*	*P. austriacus*	*R. hipposideros*
D510	ACTGTACG	*P. austriacus*	*P. austriacus*	*P. nathusii*	*P. nathusii*	*P. nathusii*	*P. nathusii*	*P. nathusii*	*P. nathusii*	*R. hipposideros*
D511	GCGCATTC	*P. pipistrellus*	*P. pipistrellus*	*P. pipistrellus*	*P. pipistrellus*	*P. pygmaeus*	*P. pygmaeus*	*P. pygmaeus*	*P. pygmaeus*	*R. hipposideros*
D512	CTCTGGAA	*P. pygmaeus*	*P. pygmaeus*	*P. pygmaeus*	*P. pygmaeus*	*R. ferrumequinum*	*R. ferrumequinum*	*R. ferrumequinum*	*R. ferrumequinum*	Blank

50 µl PCR reactions were prepared with the appropriate indexes using Amplitaq gold (Moretti et al., 1998), then subjected to the following thermal cycler settings: 95°C for 12 mins, then 10 cycles of 95°C for 20 s, 60°C for 30 s, 72°C for 40 s, followed by a final extension stage of 72°C for 5 mins. These were purified using 50 µl of homemade SPRI beads, then eluted in 20 µl of EBT.

3.3.1.4.8. PCR round two

A second 50 µl PCR was undertaken using generic indexing primers and Phusion master-mix, using the following thermal cycler settings: 98°C for 30 s, 10 cycles of 98°C for 20 s, 60°C for 30 s, 72°C for 40 s, followed by 72°C for 5 mins. These were cleaned and size selected using 50 µl of homemade SPRI beads, then eluted in 20 µl of EBT to remove remaining adaptors and indexing primers, shown in figure 3.6.

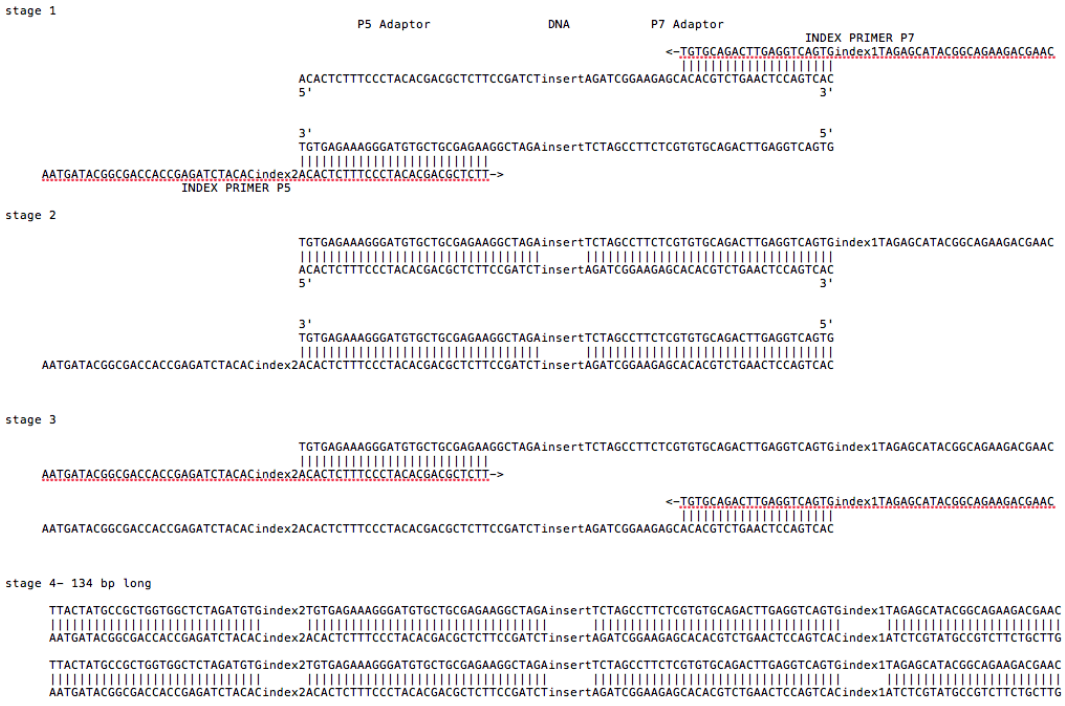


Figure 3.6: Indexing PCR process

3 μl of each sample was gel electrophoresed to establish library preparation success and determine library insert size.

3.3.1.4.9. Quantifying and pooling libraries

Final library concentration was determined using Qubit BR reagents. The molarity of each library was calculated using eqn 1, and was diluted to 4nM, with the dilution factor calculated using eqn 2, and the indexed libraries were pooled. The blank library was not diluted.

$$\text{pmol DNA} = \mu\text{g DNA} \times \left(\frac{\text{pmol}}{660\text{pg}}\right) \times \left(\frac{10^6\text{pg}}{1\mu\text{g}}\right) \times \left(\frac{1}{\text{N}}\right)$$

(Eqn 1, dsDNA length and concentration to pmol DNA)

Where 660 pg/pmol is the molecular weight of a pair of nucleotides, and N is the number of nucleotides.

$$\left(\left(\frac{\text{pmols/mL}}{4}\right) - 1\right) : 1$$

(Eqn 2, volume of H_2O to add to 1 μl sample to dilute to 4nM)

3.3.1.4.10. Running the Illumina MiSeq

The 4nM pooled libraries were denatured using freshly prepared 0.2N NaOH. This was then diluted with HT1 to create a 12pM library with 1mM NaOH. This was spiked with 12.5pM of PhiX Control. These were run on an Illumina MiSeq using a V2 2x250 (500bp) cartridge, as shown in figure 3.7.

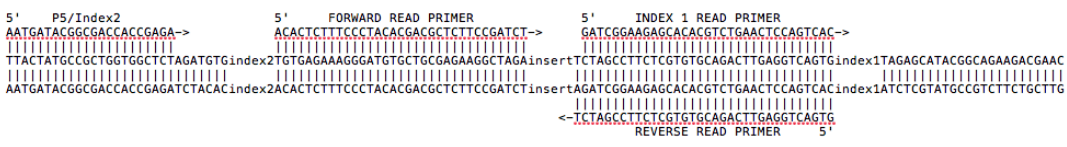


Figure 3.7: The sequencing PCR process

After sequence processing (de-multiplexing, adapter removal, filtering, artificial duplicate read removal, BLAST analysis) the two datasets were merged.

3.3.2. Bioinformatics

Except for de-multiplexing, which was undertaken on a windows machine, all bioinformatics was undertaken on either on a Macintosh computer running OSX version 10.9.5, or on an Ubuntu server, version 12 with 644GB memory, 64 x Intel Xenon E7s at 2.14 GHz each. Bash was version 3.2.53, Perl was version 5.16.2, and R scripts are written to run on 3.2.2 (R Developement Core Team, 2013).

3.3.2.1. De-multiplexing sequences

Data was removed from the MiSeq. Due to the use of custom indexes, MiSeq Reporter was used to assign raw FASTQ index reads to samples, and demultiplex them. The sample spread sheets used in MiSeq Reporter to demultiplex the samples are attached in appendix B.1.

3.3.2.2. Adapter removal and quality filtering

FASTQC was used to analyse the quality of the sequence data and the index sequences, before and after filtering. These identified a number of over-represented sequences, which on closer inspection proved to be un-trimmed Illumina sequencing adapters. As MiSeq Reporter was found to have not fully remove all adapter sequences, these were removed using CutAdapt, which was also used to remove low quality sequences (with a score threshold of 10, see below) (Martin, 2011).

```
$ cutadapt -q 10 -B <adapter sequence> -b <adapter sequence> -o
<output.1.fastq> –p <output.2.fastq> <input.1.fastq>
<input.2.fastq>
```

3.3.2.3. NCBI BLAST analysis

BLASTn was selected to compare the sequences to the “non-redundant” (nr) nucleotide database, which is the largest NCBI database available (Altschul et al., 1990). Additionally, BLASTn outputs can be parsed by MEGAN5 (see section 3.3.2.4) which is used in later analyses (Huson et al., 2007).

3.3.2.3.1. Sequence conversion from FASTQ to FASTA

Once binned into samples and quality filtered, sequences were converted from FASTQ to FASTA format using the following bash command as a model.

```
$ cat <input.fastq> | awk 'NR%4 !=0' | awk 'NR%3 !=0' | sed
's/@/>/g' ><output.fasta>
```

This formatted the sequences to be BLASTn compatible. After quality filtering, artificial duplicate reads (ADR) were removed from the dataset, as these are as a result of the PCR amplification of the library, and do not reflect the true metagenome (Gomez-Alvarez et al., 2009). ADRs were removed using fastx_collapser from the fastx toolkit (Gordon and Hannon, 2010).

```
$ fastx_collapser -i <input.fasta> -o  <collapsed.fasta> -v -Q33
```

3.3.2.3.2. Sequence length distribution

The sequence lengths for each of the libraries was summarised using the following bash code as a model.

```
$ cat <input.fasta> | awk '{if(NR%2==0){print length($0)}}' |sort
| uniq -c ><output.txt>
```

These were then combined and summarised to give a breakdown of all of the sequence lengths (figure 3.8).

3.3.2.3.3. NCBI BLAST database creation

The latest NCBI BLAST nucleotide database was downloaded onto a server. It was configured with multithreading and using 64-bit binary using the following code:

```
$./configure --with-mt --with-64
```

This specifies the use of multithreading (up to 16 threads) and 64 bit for more efficient memory usage.

Multiple executable shell scripts were created, which would allow the programs to remain with in memory allocation limits, with the following code for each of the sample files (forward read and reverse read):

```
$./blastn -db nt -num_threads 8 -query <filename> -outfmt 7 -out
<filename>.out -num_alignments 10
```

Using a tabulated output, for speed, it was possible to finish the blast run for the forward reads in under a week, and the reverse in the same amount of time.

3.3.2.4. Parsing BLAST output with MEGAN

Using MEGAN (MEtaGenome ANalyzer) command line, the blast output .txt files and the FASTA files were merged and converted into .rma files.

```
$./JavaApplicationStub +g false -x "import blastfile=<input.txt>
fastafile=<input.fasta> meganfile=<input.rma> minscore=100”
```

The command “minscore=100” ensured that low quality hits (below a BLAST threshold of 100) were categorised as “Not Assigned”. It is possible to create .rma files directly from blast outputs without using .fasta files, which is advisable where file sizes are limiting, however,

where file sizes are not limiting, this is useful as it allows easy identification of sequences of interest.

An .rma file was created by merging all of the blast files, in order to give an overview of the results. Individual .rma files were also created for each of the species by merging the all of the appropriate .rma files from within the MEGAN GUI. Sequences can be extracted from subsets of .rma files (such as all hits assigned to “Arthropoda”) using the following MEGAN5 command to call the MEGAN command:

```
$./JavaApplicationStub -g +s -E +w –c <Commandfile.txt>
```

Example MEGAN5 command from within the Commandfile.txt:

```
"open file='<infile.rma>'; extract what= reads
outdir='<outfile_location>' outfile='<outfile.txt>'
data='Taxonomy' names='Arthropoda' allbelow='true'”
```

This allows for focused comparison of samples and was used to extract the Arthropoda, Chiroptera, bacteria, fungi and virus datasets.

3.3.2.5. Verifying shotgun metagenome data

Reliable phylogenetic assignation of reads of metagenomic DNA is a non-trivial issue because of variable database representation across organism kingdoms for different genomic regions. This is particularly problematic the organisms being studied are closely related to model orgasms.

3.3.2.5.1. Phylogenetic intersection analysis (PIA)

If the genomic region of a DNA sequence from a sample is not represented by its organism species in the database, then this can result in the DNA sequence being attributed to the ‘next closest’ taxa sequence available. Therefore, there can be a tendency to assign sequences to taxa that are over represented in the database, which may make no sense ecologically and will often include model organisms. Such

sequences may be identified as misassignment, or as adapter contamination, and require post hoc removal that demands an interpretation of the data based on factors other than the information value of the genetic data alone. This is especially problematic in instances where the true species are both unexpected in the study context and model organisms. Many protocols rely on good database representation, for instance through using barcoding targets for which a large number of species have been surveyed, and assignation is dependent on a high threshold of similarity between data and database entries.

PIA is designed to be used in conjunction with BLAST (Altschul et al., 1990) and MEGAN (Huson et al., 2011).

3.3.2.5.1.1. How does PIA work?

1. File checks- the first stage of the PIA programme is checking that the input files exist, can be opened, and that the file contents are formatted correctly.
2. Get lineage- each sequence hit is assigned a full taxonomic classification up to superkingdom, using the NCBI taxonomy files.
3. PIA intercept drop-off
    a. Get the taxonomic classification of the first sequence hit
    b. Iterate through the following sequence hits
    c. When a novel classification is found, check previous line for score (bits)
    d. If score is the same as this novel sequence hit, assign this as intercept- ‘I1’
    e. Find next score change with a novel species
    f. Test to see if the intercept is the same as ‘I1’, if it does, repeat e.
    g. If it does not match, assign this as the new intercept- ‘I2’
    h. Find next novel drop off, if it does not match with “i2’, accept the ‘I2’ intercept and Assign the sequence to this classification

4. Summarise output.

3.3.2.5.1.2. Using the PIA toolkit to adjust data by genome size

PIA comes with the option of an extended output format. It also has the option of allowing for adjustment of the proportion of the reads assigned to each taxa by its genome size. Where there is no published genome size available for a particular taxon, the Genome Size Adjuster (GSA), will adjust by the closest available classification with a genome size available. At present, a spread-sheet containing information about many plant and animal taxa, however, the option to provide a custom spread sheet is available (this must be formatted as in the same way as the genome size spread sheet, appendix E.3.5). More information about the Genome Size Adjuster can be found in 3.3.2.6.

3.3.2.5.1.3. Running PIA and arguments

PIA was carried out on the Arthropoda dataset. As PIA requires a full BLAST output (rather than the much faster tabulated output used previously), BLAST was re-run on the Arthropoda dataset using the following command:

```
$ blastn -db nt -num_threads 8 -query  <fasta.file> -out
<output.file> -num_alignments 1
```

PIA.pl is designed to be able to take arguments from the command line using flags read by the getopts package. Usage is as follows:

```
$ perl PIA.pl -f <FILE> -b <FILE> [options]
```

Table 3.6: PIA main arguments

If an input is required, a Y is specified in the "input required" column, if there is an "N", no additional information is required.

Option	Description	Input required	Explanation
-f	File name	Y	Enter filename- a list of fasta headers extracted from MEGAN

-b	Blast File name	Y	Blast filename- full blast output file (not tabulated output)

Table 3.7: PIA optional arguments

If an input is required, a Y is specified in the "input required" column, if there is an "N", no additional information is required. If "optional" is specified, then an alternative file may be specified; otherwise the default will be used.

Option	Description	Input required	Explanation
-c	Cap	Y	A cap to use in the PIA analysis- determines the calculation of the taxon diversity score
-n	Path	Optional	Relative path to nodes.dmp default path is ./Reference_files/nodes.dmp
-N	Path	Optional	Relative path to names.dmp default path is. /Reference_files/names.dmp
-h	Help	N	Print help-file
-e	Extended summary	N	Option to create an extended summary file- can be slow
-g	Genome adjustment	N	Option to undertake genome size adjustment- uses default. /Reference_files/All_Genomes_SS.txt
-G	Genome Spread-sheet	Optional	Option to provide alternative genome size spread-sheet name or location

PIA.pl uses the taxonomy tree information from the NCBI files nodes.dmp and names.dmp. The latest versions can be found here: ftp://ftp.ncbi.nih.gov/pub/taxonomy/taxdump.tar.gz

3.3.2.5.2. Assessing the sensitivity of shotgun metagenomics relative to metabarcoding

The proportion of mitochondrial reads in a shotgun metagenome dataset can be used as a measure of the sensitivity of the shotgun metagenome dataset (Smith et al., 2015). On the *Manduca* sexta (the tobacco hornworm) mitochondrial genome, there are 13 protein coding genes, which comprise 72% of the mitochondrial genome (11,185bp out of the 15,516 bp mt genome) (Cameron and Whiting, 2008). The metabarcoding

loci used in this and previous studies (chapter 4) (Zeale et al., 2011), comprises 1.4% of the mt protein coding regions. The number of Arthropoda hits for each sample was collected (appendix E.3.4), and the number of protein mitochondrial Arthropoda BLAST hits was determined using the following command:

```
$ cat <arthropoda_blast_file.txt> | grep "^>" | grep ' COI \| cytB \| cytochrome \| cox1 \| cox2 \| atp8 \| atp6 \| cox3 \| nad3 \| nad5 \| nad4 \| nad4L \| nad6 \| cob \| nad1' >> <outputfile>
```

The number of protein mitochondrial Arthropoda BLAST hits were then counted (appendix E.3.4). We assume that all of the mt protein coding regions were assigned in the dataset, and were extracted and counted using the above command, the data was then adjusted by the size of the protein coding region compared to the full mitochondrial genome.

3.3.2.6. Genome size adjuster

This is a program to take the number of hits to each taxon from a metagenomic dataset and adjust them according to the genome size of the organism. Genome_Size_Adjuster.pl is designed to take the output from MEGAN or PIA. Where the genome size of the organism is unknown, the program moves up the taxonomic levels until a genome size is found. Where there is more than one genome size available for a particular taxon (which is especially common at the higher taxonomic levels), an average of the available genome sizes is used.

3.3.2.6.1. Running Genome_Size_Adjuster.pl

Genome_Size_Adjuster.pl uses the taxonomy tree information from the NCBI files nodes.dmp and names.dmp. The latest versions can be found here: ftp://ftp.ncbi.nih.gov/pub/taxonomy/taxdump.tar.gz. Usage is as follows:

```
$ perl Genome_Size_Adjuster.pl -f <FILE> -g <FILE> -n <file> -N <file> [options]
```

Table 3.8: Genome size adjuster main arguments

If an input is required, a Y is specified in the "input required" column, if there is an "N", no additional information is required.

Option	Description	Input required	Explanation
-f	File name	Y	Enter filename- an extraction from MEGAN of no. Hits assigned to each taxa
-g	Genome file name	Y	Genome size database
-n	Path	Y	Relative path to nodes.dmp
-N	Path	Y	Relative path to names.dmp
-h	Help	N	Print help file

Chiroptera and Arthropoda genome sizes were collated (amongst others), and can be found in appendix E.3.1., E.3.2., E.3.3.

Before undertaking the size adjustment, the program filters the data in two ways. Firstly, any rows containing zero only values are removed (these are quite common in MEGAN outputs due to the inclusion of intermediate nodes). Secondly, if any names are formatted incorrectly, or do not appear in the names.dmp file, are removed from the dataset. This is printed to the screen and also to an error file.

The Genome Size Adjuster produces two key spread sheets, one showing the original data (original.txt), and one with the adjusted data (adjusted.txt). A number of intermediary spread sheets are also created.

After the data have been adjusted, Genome_Size_Adjuster.pl calls an R script to create a figure that shows the before and after hit values as percentages in a stacked bar chart. These are clustered using the R hclust distance complete linkage clustering method (Oksanen, 2010).

In order to run the output of PIA in Genome_Size_Adjuster.pl, the script PIA_to_GSA.pl is used. This will put the data into the appropriate format (as above). If Genome_Size_Adjuster is called from within PIA, this formatting is done automatically. Both of these scripts can be found in appendix D.

3.3.2.6.2. Known bugs

Data must not be from taxa higher than phylum (i.e. only select nodes lower than the phylum of interest). If hits are assigned to kingdom are used, it will most likely crash the program.

3.3.3. Statistical analyses

3.3.3.1. Rarefaction analyses

Rarefaction data for the full dataset, a subset of just the Chiroptera data, and a subset of just the Arthropoda data, was extracted from MEGAN5. These were then plotted using R (code in appendix D) (Chroňáková et al., 2009, Heck Jr et al., 1975). Rarefaction plots are constructed from the expected number of species E(S) from a sample of a smaller size. The expected number is calculated using equation 3.

$$\mathrm{E(S)} = \sum_{i-1}^{S} \left(1 - \left[\frac{\left(\frac{N - N_i}{n} \right)}{\left(\frac{N}{n} \right)} \right] \right)$$

(Eqn 3, Expected number of species),

Where N is the number of individuals in the rarefied sample, N_i is the number of individuals in the i^{th} species, and n is the size of the subsample.

3.3.3.2. Calculating dietary diversity and niche breadth

Dietary diversity and niche breadth were calculated as in 2.3.2.1.

3.3.3.3. Determining trophic niche overlap

Trophic niche overlap was calculated as in 2.3.2.2.

3.4. Results

The first Illumina run (which had clustering problems) returned 1.241 GB of FASTA data after quality filtering, of which 119 MB was undetermined: during the de-multiplexing process, these data were not assigned to any sample. This happens as a result of errors in the index sequences introduced during PCR amplification, due to errors in either of the indexing read cycles of the MiSeq run or in the base calling stages. The second Illumina run returned 8.46 GB of data, of which 515 MB was classified as undetermined. The second Illumina run returned a large amount of data for a MiSeq run (which advertises a yield of up to 8.5 GB) (all sequences with a quality score (Q score) of >30).

3.4.1. Hits before and after adapter removal and quality filtering

Initial, unfiltered BLAST searches returned some unexpected results; there were a number of hits reported for wild Bactrian camel (*Camelus ferus*) and for Tasmanian devil (*Sarcophilus harrisii*). This is likely an artefact of poor adaptor trimming in the *C. ferus* and *S. harrisii* data. However, after re-doing adaptor trimming using cutadapt, doing quality score trimming, and filtering all hits which had a BLAST score <100, these hits were both entirely removed from the dataset. Crucially, the data that we were expecting to find (chiropteran and arthropod) was not significantly reduced by the filtering process. Furthermore, when a BLAST database was created of the Illumina TruSeq adaptors, *S. harrisii* was found to be a good match. After filtering, the datasets were merged giving a final 9.7016 GB dataset. FastQC outputs for both datasets can be found in appendix C.3.2.

3.4.2. DNA fragmentation and damage

Agilent 2100 Bioanalyser high sensitivity DNA (Panaro et al., 2000) traces of the raw DNA extractions had a mean sequence length distribution of >250 bases. The library sizes after SPRI bead size selection, as

determined using gel electrophoresis, were between 500 and 700bp. The lengths of sequenced data can be seen in figure 3.8; the majority of the sequences are 250 bp long; the maximum read length of the Illumina MiSeq.

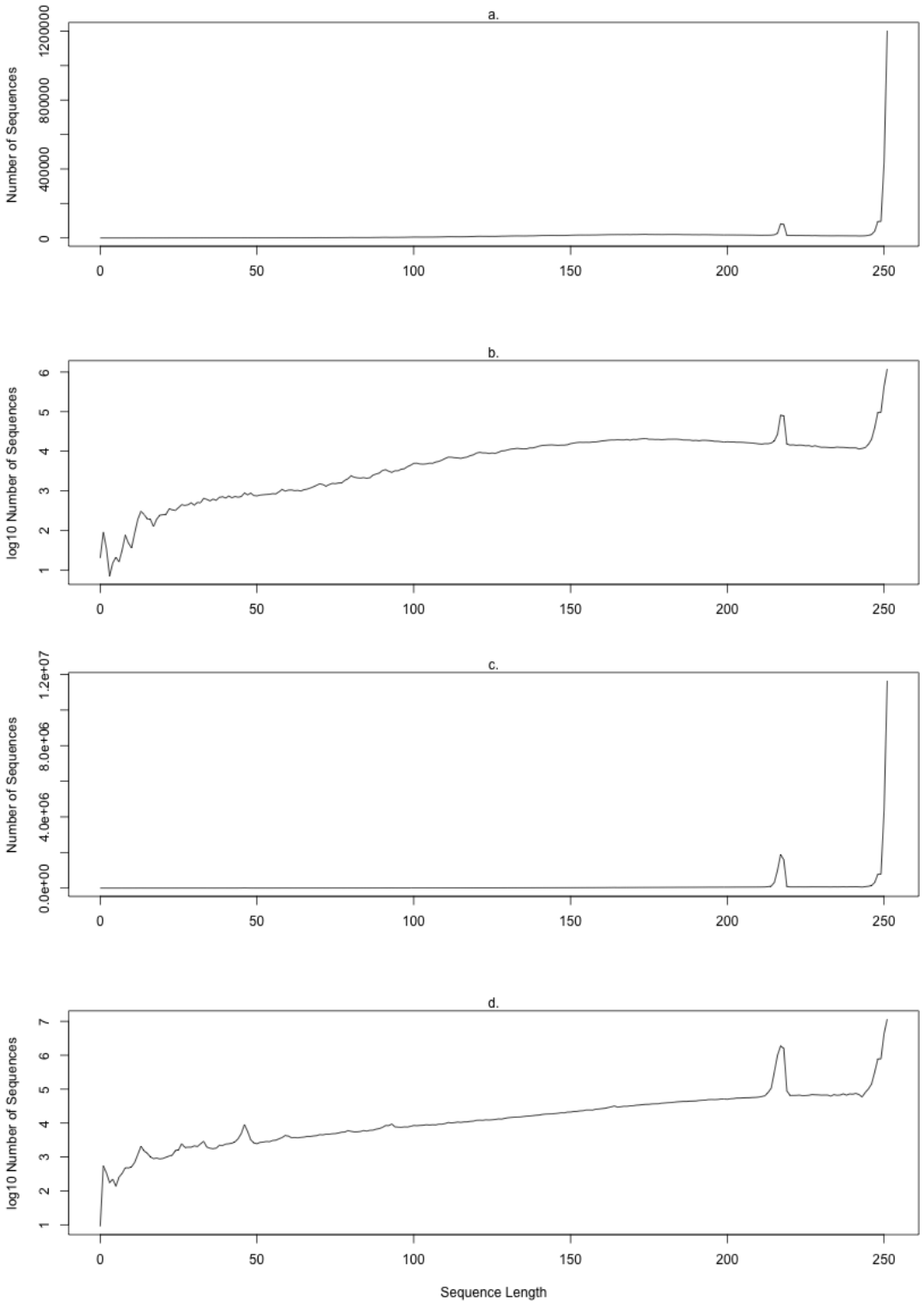


Figure 3.8: Fragment length distributions

From the first Illumina run (A and B) and the second Illumina run (C and D). A+C. Sequence length against the number of species, B+D. sequence length against the log10 number of sequences. Code in appendix D.2.6.

The fragmentation process was effective; the clustering of sequences on the MiSeq was high, with high quality reads, suggesting that there were no issues with DNA strands that were too long interfering with cluster formation. At the same time, it does not fragment the DNA too far; the long sequence read length implies that over fragmentation did not occur.

Due to the long read length (250bp) of most sequences, and the metagenomic nature of the data, damage analysis programmes such as mapDamage (Ginolhac et al., 2011) are not suitable for use on these data.

3.4.3. De-multiplexing sequence data

Table 3.9 shows the data assigned to each of the bat species, and the average data per library. *P. austriacus* returned the largest dataset, at 734.13 MB of data. *M. alcathoe* returned the highest data per library at an average 71.04 MB per library, and *M. brandtii* had the lowest data returned per library at 35.20 MB per library. The average FASTA data yield per library from the first Illumina MiSeq run was 7.91 MB. The average yield per library from the second MiSeq run was 38.20 MB. Table 3.9 also shows the data assigned to each bat species and to the controls. The first Illumina run contained an index combination for which no library was created (table 3.4, indexes D709 and D508). There were no reads assigned to this combination after demultiplexing, suggesting that there is no observable chimeric misassignment of index reads.

Table 3.9: Breakdown of the data returned for each species, and data assignment
This table is summarised from appendix E.3.4. These data are after adapter removal, quality trimming and blast score filtering.

Species	Number of libraries	Final Number of	Total MB of filtered FASTA	Number of Total Hits	Number of Chiropteran BLAST hits	Number of Arthropod

		samples	Data			BLAST hits
B. barbastellus	10	6	451.32	357,218	3497	3597
Blank	2	1	39.96	21,281	35	21
E. serotinus	10	7	416.94	319,765	1717	824
M. alcathoe	1	1	71.04	35,421	9258	36
M. bechsteinii	9	6	491.03	502,058	2876	355
M. brandtii	11	7	387.15	229,876	13989	2301
M. daubentonii	10	7	573.16	475,526	13415	1879
M. mystacinus	11	7	706.26	311,899	5042	981
M. nattereri	12	7	671.76	410,644	37916	1624
N. leisleri	10	6	619.05	401,060	33450	680
N. noctula	9	6	552.64	384,903	25273	1640
P. auritus	13	7	604.8	341,455	25764	1295
P. austriacus	11	7	734.13	508,052	14713	4001
P. nathusii	10	6	526.78	371,360	6124	1577
P. pipistrellus	13	7	610.83	513,725	38479	2314
P. pygmaeus	12	7	618.83	506,769	37716	498
Positive control	2	1	19.1	30,453	2959	23
R. ferrumequinum	10	6	702.05	449,498	13489	1921
R. hipposideros	11	7	615.51	240,500	11225	1015

Generally, the bat species with larger datasets had a larger number of sequences assigned to both Arthropoda and Chiroptera (figure 3.9). However, this is not a strong trend- the R^2 for the amount of data returned against the number of Chiroptera BLAST hits (figure 3.9.a) was only 0.1857, and the R^2 for the amount of data returned against the number of Arthropoda BLAST hits (Figure 3.9.b) was only 0.2084. The points with the lowest data returned (far left in both plots) are the positive control and the blank respectively. Despite all of the species having similar numbers of samples processed, there was great variability in the number of reads returned for each species. As can be seen in table 3.9, the number of Chiropteran or Arthropod hits roughly increases with an increasing data return. However, here are a number of notable exceptions, such as *M. bechsteinii*, which, despite a large data return, yields fewer Chiroptera and Arthropoda reads than might be expected.

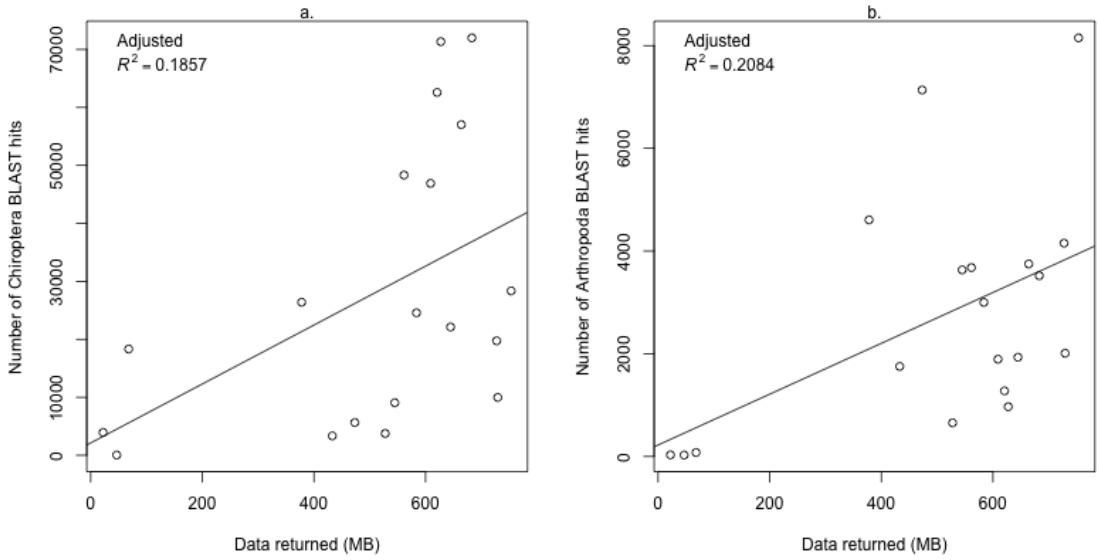


Figure 3.9: The amount of data (MB) returned for each of the species datasets Against the number of a. Chiroptera BLAST hits and b. Arthropoda BLAST hits. Code can be found in appendix D.2.7.

3.4.4. Metagenomic assignments

The majority (73.8%) of sequences with a match on the NCBI nucleotide database are assigned as bacterial in origin, which can be seen in figure 3.10. The bacterial sources may include gut microflora, environmental microflora, or bacterial DNA originating from the reagents (Salter et al., 2014a, Salter et al., 2014b). The bacterial portion of this dataset is discussed in section 3.5.3.1. The second most represented taxonomic group (10.1%) is the Chiroptera (see section 3.5.3.2), which would likely be from epithelial cells shed by the bat into the faecal matter (Eggert et al., 2005). Fungi comprise 7.95% of the assigned data. Fungi may originate from direct contact with the bat, or from the environment from which the guano had been collected (Lorch et al., 2012). Fungi is unlikely to have a dietary origin, unless prey that was infected by fungi had been consumed by the bat, this is discussed further in section 3.5.3.4. The "other vertebrates" group includes all vertebrates except the Chiroptera, and is mainly comprised of primates and rodents, which may be as a result of environmental contamination or from contamination of laboratory reagents. Additionally, it may be as a result of over-representation of these taxa on the NCBI database, and subsequent misassignment of

conserved DNA sequences to these taxa. Arthropoda DNA, which comprises 1.01% of the assigned dataset, is likely from dietary sources, and is discussed in sections 3.5.4 to 3.5.7. Viruses (0.533%) are discussed in section 3.5.3.5, are mostly dsDNA viruses, and the majority of the virus sequences come from two *P. pipistrellus* samples. Other DNA sources include the Viridiplantae, protist, algae, and nematode.

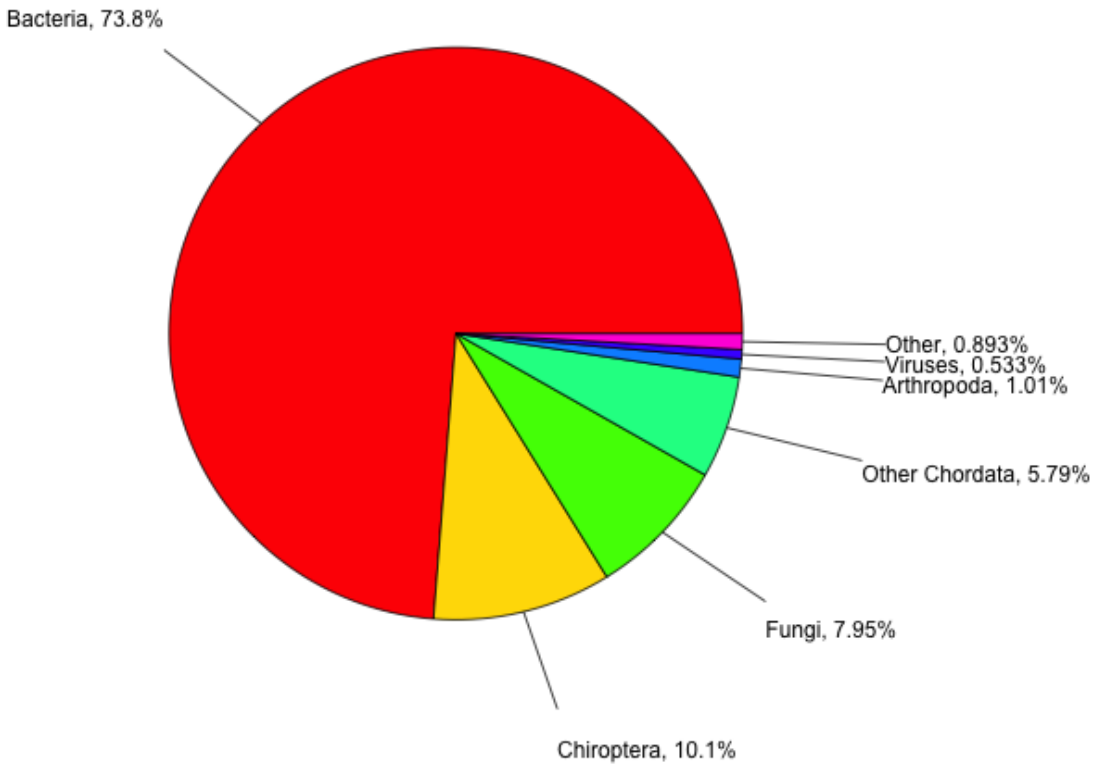


Figure 3.10: Metagenomic overview of Illumina shotgun sequencing data for all samples

Code in appendix D.2.8.

3.4.4.1. Negative control

The blank libraries returned 39.96 MB of data, with a total of 21,281 sequences, of which 3304 were assigned. The average data returned per library was slightly more than that returned from the positive control library (table 3.9) although it was much lower than the average amount of data returned for the sample libraries, and is probably more successful than some of the other libraries due to the lack of PCR inhibitors, which commonly occur in guano. Additionally, where the other libraries were

diluted to 4nM before sequencing, the blank libraries were not diluted. Figure 3.11 shows the breakdown of the sequences returned. There were very few bat sequences returned, indicating that there has been no contamination of the reagents by the samples. This suggests that all of the sequences returned from the blank library originated from the reagents (Salter et al., 2014a, Salter et al., 2014b).

There was a small proportion of Arthropoda reads in the blank libraries (35 reads), all of which were assigned to Diptera. All of these reads originated from the blank library from the first Illumina MiSeq run. 12 of these reads were assigned to the Tipulidae, and thus data from the first Illumina run, that is assigned to Tipulidae, should be discounted, unless the proportion of the reads assigned to Diptera is significantly higher than in the blank library.

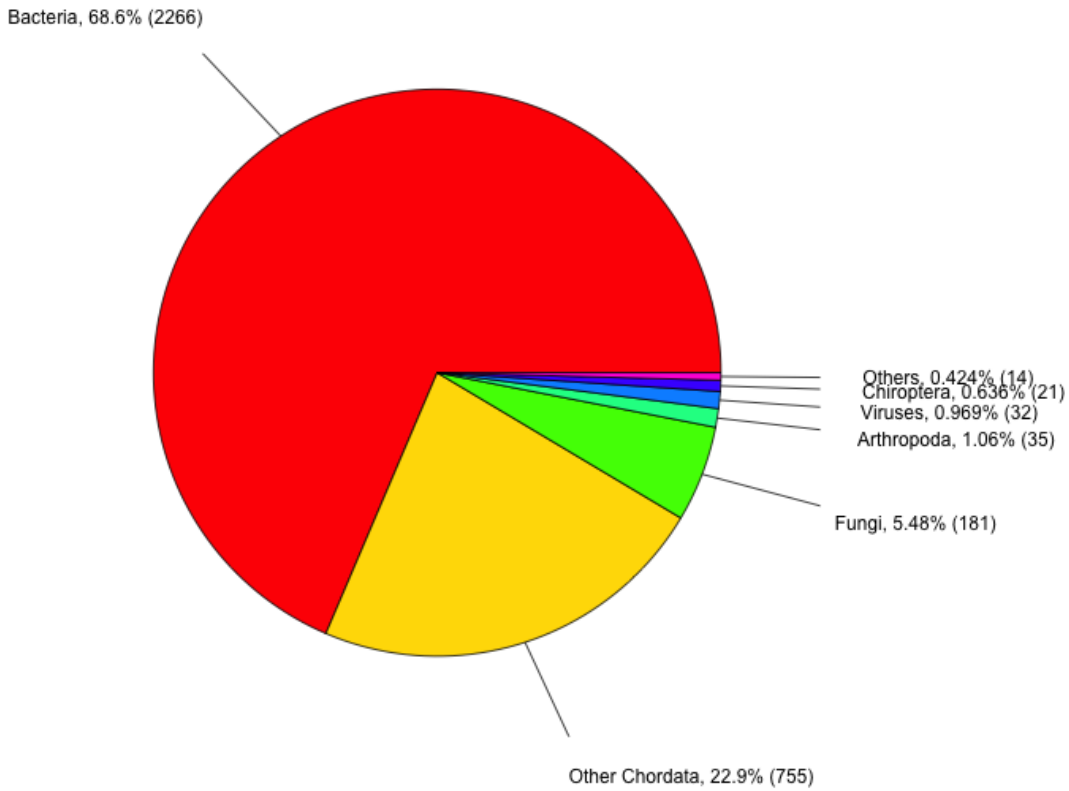


Figure 3.11: The taxonomic profile of both the blank libraries combined
Showing the number of sequence hits in brackets. Code in appendix D.2.8.

3.4.4.2. Positive control

As a positive control, a library was prepared from a captive *P. pygmaeus*. This returned only a small number of arthropod reads (23 reads), which were mostly identified as being dipteran in origin. This was surprising, as the bat had been fed on *Tenebrio molitor*, which is a coleopteran. On inspection, the majority of the hits were to *Drosophila* species, which is unsurprising, as this genus contains model organisms, and thus composes the majority of the Arthropoda databases. The barcoding stage of this project will likely prove more informative as the database for COI barcodes is better developed and phylogenetic analysis will be possible on the sequence data. However, it must be remembered that despite being hand fed *T. molitor*, it is not impossible that the captive bat was also self-feeding. However, it is also possible that these DNA were contamination, as all of the Arthropoda reads in the positive control libraries came from the positive control library from the first MiSeq run, which had similar (low) proportion of Dipteran contamination. It is possible that no arthropod DNA from this sample was sequenced, and that these Diptera reads were contamination (section 3.4.4.1).

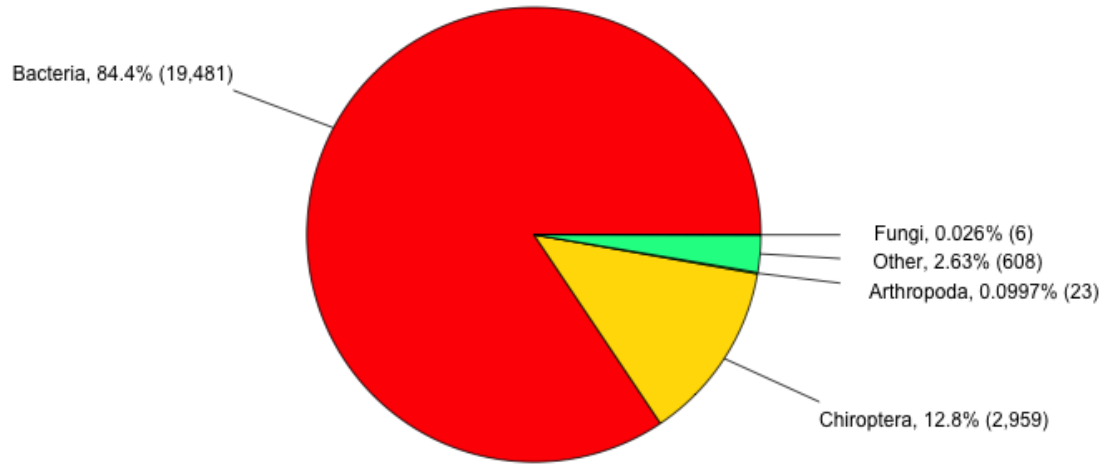


Figure 3.12: The taxonomic profile of both the positive control libraries combined
The positive control was a captive *P. pygmaeus.* Showing the number of sequence hits in brackets. Code in appendix D.2.8.

3.4.5. Rarefaction analysis

Figure 3.13. shows the rarefaction plot for the full dataset at order level, and shows that all plateau quickly. The alcathoe, positive, and negative control datasets have a far lower number of leaves (Taxa assigned at order level), which is probably an artefact of the lack of technical replicates. This suggests that the sampling effort has been sufficient for identifying the majority of the taxa present across the full dataset (Chroňáková et al., 2009, Heck Jr et al., 1975).

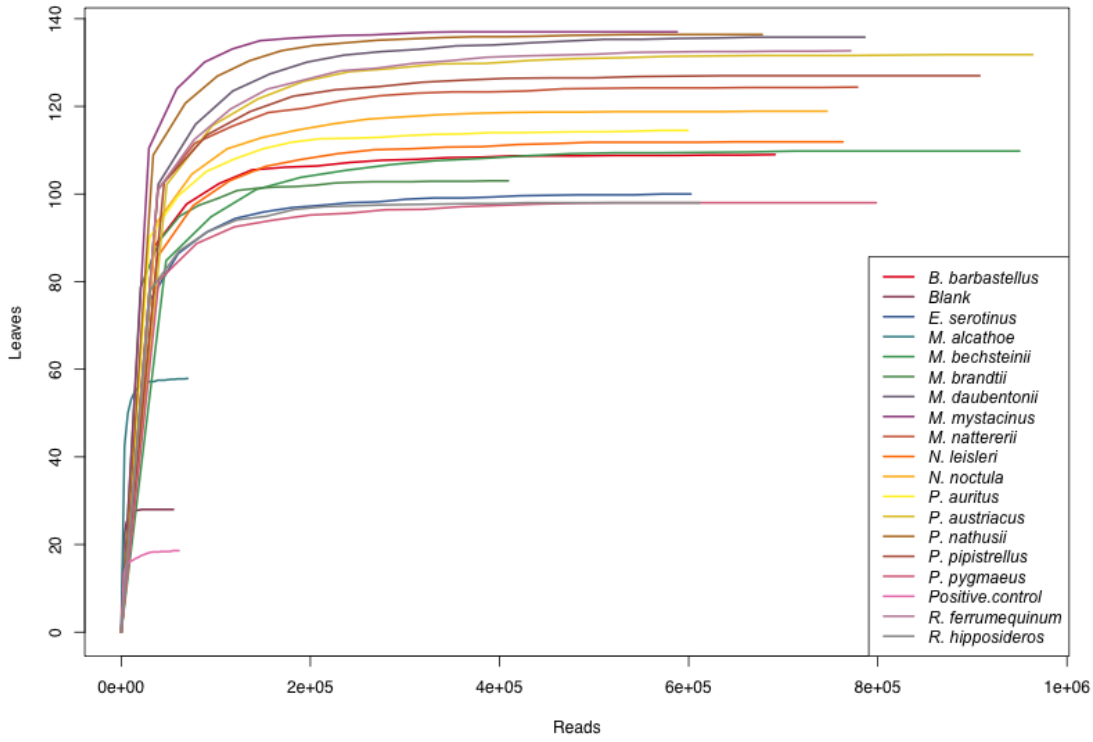


Figure 3.13: Rarefaction plot for the full dataset at order level

Including controls. “Leaves” are number of taxa at order level. Code in D.2.9.

Rarefaction plots for the Chiroptera and Arthropoda data subsets can be found in appendix C.3.3.

3.4.6. DNA sources

3.4.6.1 Bacterial

21 bacterial phyla were identified in the full dataset, with 5 of these common across the guano of all of the studied bat species. The high

proportion (73.8%) of sequences returning as bacterium is unsurprising. It will likely consist mainly of bat gut microflora, and of environmental bacterium transferred to the sample upon deposition of the guano, bacterium introduced during sampling (and library preparation), and from the reagents (see section 3.4.4.1).

Figure 3.14. shows the breakdown of the bacteria identified as associated with each bat species; which shows the metagenomic assignments collapsed at phylum level. The top four phyla represented in the full dataset are Proteobacteria, Actinobacteria, Firmicutes, and Bacteroidetes, all of which are large constituents of gut microflora in humans (Barka et al., 2016) and in the guano of bats (De Mandal et al., 2015). In all species, the majority of sequences (between 60.22% and 90.79%) are assigned to the Proteobacteria. The Proteobacteria are a large, diverse phylum of Gram-negative bacteria which include a large number of animal symbionts and pathogens (Ramulu et al., 2014). The high proportion of Proteobacteria differs to the previous studies on bats (De Mandal et al., 2015), and is higher than is typically seen in humans, which have low proportions of faecal Proteobacteria, compared to that seen in this study (Rajilić‐Stojanović et al., 2007).

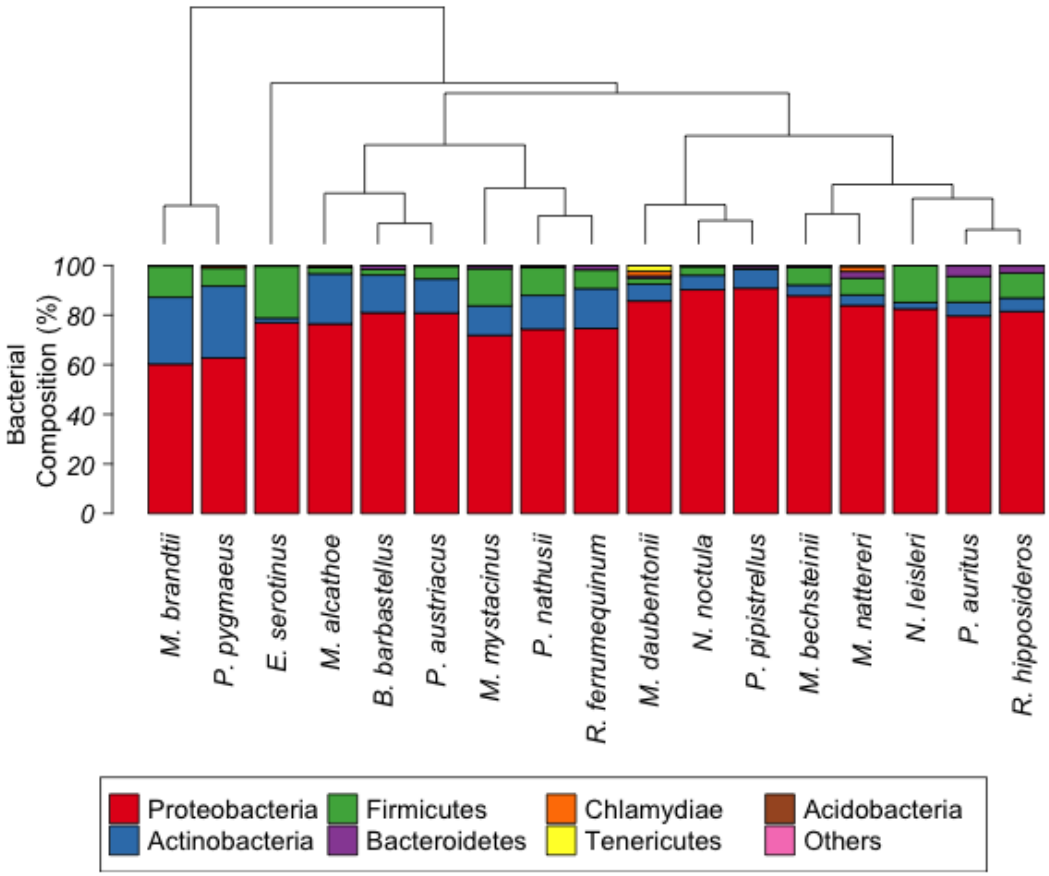


Figure 3.14: The bacteria associated with each bat species at phylum level
Prey taxa have been grouped by order. Data sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2.

The Proteobacteria are the highest represented phylum in each of the species datasets. In order to determine if this high representation was due to disparities in the genome size, we investigated whether the genome size of the Proteobacteria was larger than that of the other phyla seen. Using the data from the prokaryote table of the genome size database, the average genome size for the Proteobacteria was 3528kb (Islas et al., 2004). In comparison, the average genome size for all of the Prokaryotes listed was 3214kb. As a result, we conclude that the high representation of the Proteobacteria is largely representative of the breakdown of the bacterial phyla in the samples, and not an artefact of genome sizes. A number of the Proteobacteria are pathogenic (Nakagawa et al., 2007), discussed in section 3.4.6.1.2.

The second largest phylum represented (figure 3.14) is the Actinobacteria. Of the five subclasses of Actinobacteria (Stackebrandt et al., 1997), the Actinobactedidae is the best represented subclass, and the Actinomycetales is the highest represented order in these data. Actinobacteria are Gram-positive bacteria which can be environmental inhabitants, animal pathogens, or gastro intestinal commensals (Barka et al., 2016). *Actinobacteria* are particularly successful in the acidic conditions of bat guano (De Mandal et al., 2015, Goodfellow and Williams, 1983).

The Firmicutes is the third largest phylum in the full dataset (figure 3.14). For most of the bat species, the majority of reads map to the Bacilli, which are obligate or facultative aerobes, except in *E. serotinus*, where the majority of the Firmicutes reads map to the anaerobic Clostridia (discussed in section 3.4.6.1.1 and 6.4.6.1.2). The ratio between the Firmicutes and the Bacteroidetes is discussed in section 3.5.4.1.

Chloroflexi is surprisingly underrepresented in the dataset, only assigned 111 sequences across the full dataset, compared to 29.97% of the guano bacteria found in (De Mandal et al., 2015). The Chloroflexi are common inhabitants of cave microflora (Barton et al., 2014), which was where De Mandal's samples were collected, suggesting that they may inappropriately assigned these DNA as originating from guano, rather than from the environmental context. Alternatively, it is possible that the Chloroflexi were over represented in their dataset due to the introduction of PCR biases inherent in Metabarcoding. Finally, the bat species from which the samples originated (this is not detailed in the paper) may have different gut microflora to the Great British bat species.

78,516 bacterial reads were submitted for phylogenetic intersection analysis (up to 500 reads per sample). Of these, 24,917 were assigned after analysis (31.73% retained). After PIA, the Proteobacteria were still the highest represented phyla making up an average of 67.54% of the bacterial reads, which was followed by 14.88% assigned to

Actinobacteria. Firmicutes are collapsed into the Bacillus/Clostridium group (13.23%), and Bacteroidetes are the fourth largest phyla (1.28%). This provides robust support for the assignment of the bacteria.

3.4.6.1.1. Chitinase producing bacteria

Chitinases were identified to be present in bats in 1961 (Jeuniaux, 1961), with chytinolytic enzymes found in the gastric mucosa and intestine of *R. ferrumequinum*. Due to the large proportion of arthropod biomass which is comprised of chitin (Moussian et al., 2005), chininases play in important role in supplying carbon, energy, and nitrogen to the bats which is thought to be key in supporting bats, particularly in winter, when stored fat reserves are limited (Emerson and Roark, 2007, Whitaker et al., 2004). Chitinases in the bat gut usually originate from extracellularly secreted chitinases from bacteria (Whitaker et al., 2004). Chitinase producing bacteria, and their representation in the bat diets can be seen in table 3.10.

Table 3.10: The chitinase producing bacteria and their representation in the bat metagenome datasets

Chitinase producing identified using (Brzezinska et al., 2014, Cody, 1989, Kuddus and Ahmad, 2013).

Chitinase producing bacteria	**Seen in most or all bat datasets**	**Seen in low numbers (<1,000 reads) in some datasets**	**Not seen in any dataset**
Aeromonas	✓		
Arthrobacter	✓		
Bacillus	✓		
Chromobacterium		✓	
Clostridia	✓		
Cytophaga		✓	
Enterobacter	✓		
Erwina			✓
Flavobacterium	✓		
Myxobacter			✓
Pseudomonas	✓		

Serratia	✓		
Streptomyces	✓		
Vibrio		✓	

Aeromonas, a genus of the Gammaproteobacteria, in particular *Aeromonas veronii* which produces high levels of chitinases (Li et al., 2011), are highly represented in the datasets of two species: *M. daubentonii* and *N. leisleri*. Clostridium, a genus of the Firmicutes, is highly represented in *E. serotinus,* with 7,636 reads in this dataset mapping to *Clostridium perfringens*. The majority of these sequences (7,377) coming from one library (BatID 2297), obtained from a sample collected in the east midlands and submitted to the Ecowarwicker ecological forensics service in the summer of 2013. Enterobacter, another chitin producing genera of Gammaproteobacteria (Whitaker et al., 2004), are seen in most of the bat datasets, with *Enterobacter cloacae* reads mostly coming from a *N. leisleri* sample (BatID 2521), and *Enterobacter* sp. 638 from a *P. austriacus* sample (BatID 3152); which was collected in the Isle of White (off the south coast of England). Pseudomonas are the best represented in all bat species, particularly in one sample of *P. austriacus* collected from Dorset (BatID 1537), *P. putida* was found with 15,607 reads. Serratia was highly represented in one *M. bechsteinii* library (BatID 543).

3.4.6.1.2. Pathogenic bacteria

Of the bacterial pathogens reported in the literature (section 3.2.3) as having been identified in bat guano, three pathogens were reported in high levels in one or more library: *Clostridium perfringens* (See section 3.4.6.1.1) *(Hajkova and Pikula, 2007),* and *Salmonella* spp. (Arata et al., 1968, Reyes et al., 2011) which were both found in libraries made from *E. serotinus* guano, from the east midlands of England (Bat ID 2297) and from the south west of England (BatID 2540), and. Additionally, *Yersinia enterocolitica* was identified in a library from a *R. hipposideros* sample from the west midlands (Bat ID 1110).

A number of the reported pathogens (section 3.2.3) were present in low numbers of sequences (<1000 reads) in a number of the bat species, including: *Campylobacter* which had a small number of reads, although none mapped to *C. jejuni* (Hazeleger et al., 2011), *Listeria* spp. (Rozalska et al., 1998), small number of hits in all except *M. alcathoe* and *M. nattereri,* which have no hits, *Shigella flexneri* (Arata et al., 1968, Rozalska et al., 1998), which had a low numbers of hits in most bat species gut microflora, and *Bartonella* spp. (Concannon et al., 2005, Kosoy et al., 2010, Lin et al., 2012).

The following were not found at all: *Clostridium sordellii* (Mühldorfer et al., 2011a), *Borrelia* spp. (Hanson, 1970, Petney et al., 2000),*Grahamella* spp. (Sebek, 1974), *Neorickettsia risticii (Gibson et al., 2005), Leptospira* (Fennestad and Borg-Petersen, 1972, Matthias et al., 2005), and *Vibrio* spp. (Mühldorfer et al., 2011a).

3.4.6.2 Chiroptera

10.1% of the full metagenomic dataset were assigned to the Chiroptera. All of the libraries had reads mapping to multiple bat species and genera. Figure 3.15.a. shows the assignment of the Chiroptera reads from each of the bat libraries. The complete-linkage clustering puts the datasets together by their similarity, and the result has a strong reflection of the phylogenetic structure. Except for in the Rhinolophid bats datasets, the majority of the reads have been assigned to the Myotis. This is likely due to their over-representation on the Genbank nt databases (appendix C).

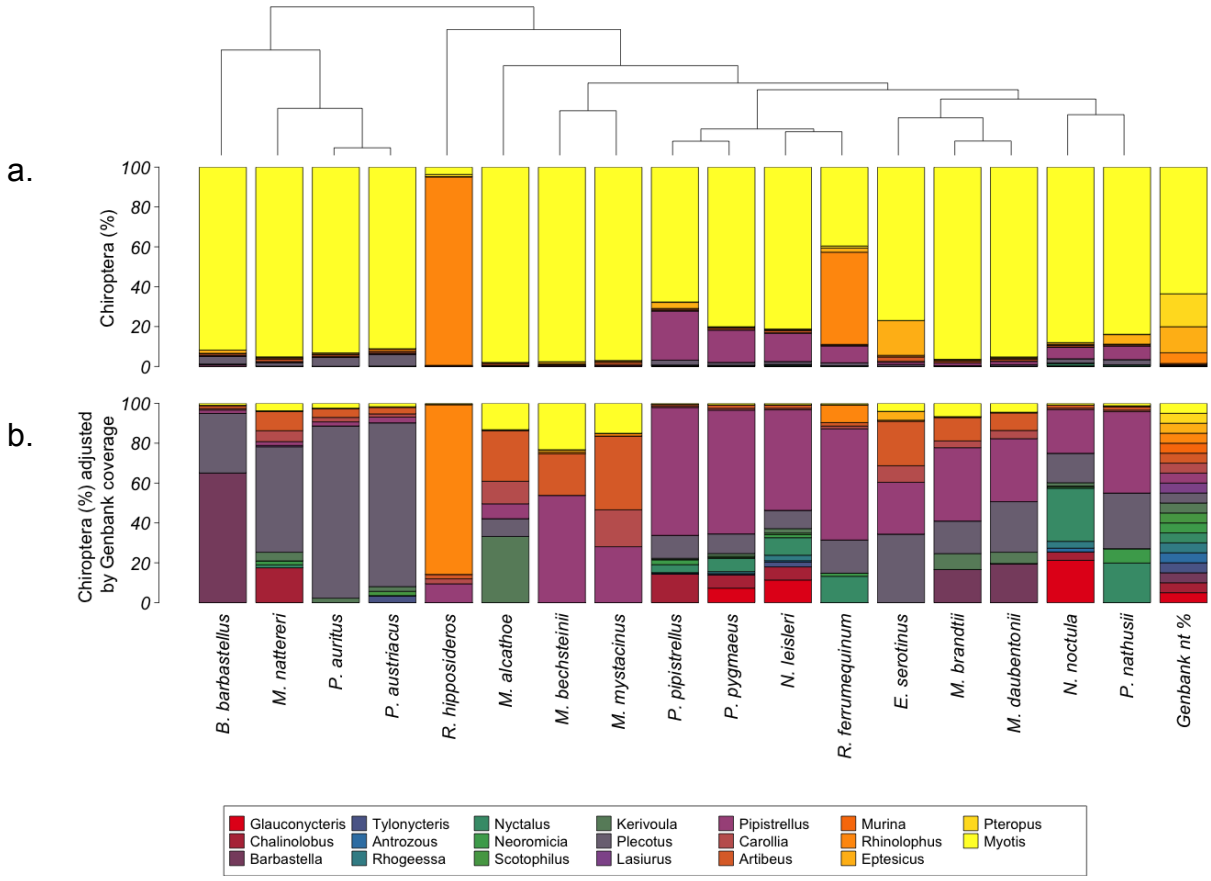


Figure 3.15: Assignment of all of the Chiroptera reads from each of the bat libraries

Data sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2. a. unadjusted data, b. adjusted by proportion of GenBank reads (which can be seen in the last column).

Figure 3.15.b. shows the same data, adjusted by the proportion of the Genbank nt entries for each genus. This dramatically reduces the proportion of reads assigned to *Myotis* in all datasets (although the *Myotis* species still have the highest proportion of *Myotis* reported). The resulting assignments do sometimes make more biological sense; the hits from the *B. Barbastella* dataset are largely assigned to *Barbastella,* both of the *Plecotus* species now have the majority of reads assigned to *Plecotus*, and the same trend is seen for *E. serotinus.* However, some species, such as *R. ferrumequinum,* are falsely identified as *P. pipistrellus.* The disadvantage of this approach is that highly represented genera (such as *Myotis*) are heavily penalised, and poorly represented genera (such as *Artibeus* and *Kerivula*) are overrepresented. An alternative approach for

such analysis would be to use a non-redundant database (such as RefSeq (Pruitt et al., 2007)) to prevent over penalization of over represented taxa such as Myotis. However, at present, RefSeq does not have all of the Chiroptera genera represented, so cannot be used here.

The proportion of Chiroptera reads was variable. It was particularly high in the positive control, which may be due to the freshness of the sample. However, some samples had a low endogenous copy number, compared to bacterial read assignment (table 3.9).

There may be many reasons for this variability; this could be due to the stochastic nature of shotgun metagenomic sequencing, may be reflective of the patchy nature of sequence databases such as BLAST, could be due to the variability in sample quality, or could be due to physiological differences in the bat species and in their digestion processes. DNA degradation may occur at different rates in the different species; digestive transport time may vary, or the digestion processes may vary, for example, in differences in enzymatic processes.

Typically for each sample, the species that the sample came from is the best represented in terms of sequence assignment. However, for all samples, a number of sequences were misassigned to a different bat species. This is unsurprising as the database is patchy, with some bat species far better represented, which is shown in table 3.6. For example, *M. brandtii* has 386,614 nucleotide records on the NCBI database, whereas *P. nathusii* only has 57 nucleotide records in total. Additionally, hits may be assigned to incorrect species due to sequences being identical between the two species. These factors are applicable to all of the data, but are most apparent within the chiropteran species. Another potential reason for hits from multiple species is where species co-habiting and depositing guanos, thus potentially causing cross contamination. All extractions were undertaken on a single guano, and all returned a clean sequence ID when subjected to PCR and Sanger

sequencing, however, the sensitivity of massively parallel sequencing is such that mixed signals may be detected.

M. brandtii has a far higher representation on the nt database than the other species (See appendix c). This is because it is the only Great British species which has had its full genome sequenced (Fang et al., 2015, Seim et al., 2013). *R. ferrumequinum* has had the mitochondrial genome of several of its subspecies sequenced, *R. ferrumequinum quelpartis* (Yoon et al., 2013), and *R. ferrumequinum korai* (Yoon et al., 2011).

This genome size adjustment was done with Chiroptera data only as this is the only data that for we know which species the data should be assigned to.

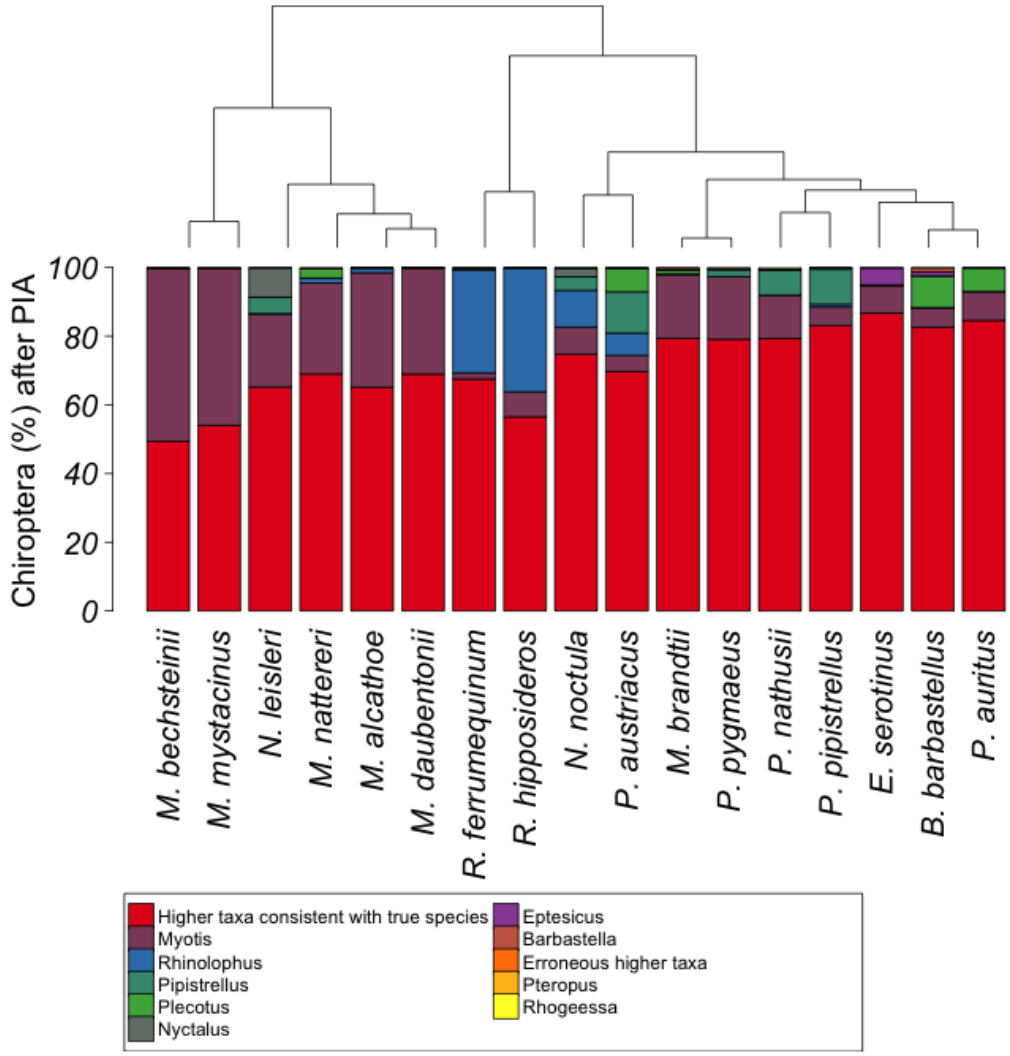


Figure 3.16: Assignment of the Chiroptera reads after PIA

Data sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2. Higher taxa include any appropriately assigned higher level of classification above the level of order and includes Boreoeutheria, Carnivora, Chiroptera, Eutheria, Lauasiatheria, Microchiroptera, Stenodermatinae, Theria, and Vespertilionidae (red). Inappropriate taxa (orange) includes Phyllostomidae, Daubenton's bat, and Perissodactyla.

36,489 Chiroptera reads were submitted for Phylogenetic intersection analysis (up to 500 reads per sample). 6,504 were assigned after analysis (17.82%). Of the reads that were accepted after PIA, the majority (71.48%) were assigned at taxonomic levels higher than order that are consistent with the true species studied (including Boreoeutheria (7.31%), Chiroptera (2.13%), Eutheria (1.17%), Lauasiatheria (7.96%), Microchiroptera (23.85%) and Vespertilionidae (29.00%), figure 3.16,

red). After PIA, the proportions of correct assignments at genus level were increased from 46.97% from the raw assignments (figure 3.15) to 58.95% (figure 3.16, table 3.11). In total, 90.45% of the post PIA assignments were correctly assigned (at genus or higher taxonomic level).

Table 3.11: The proportion of Chiroptera reads correctly assigned.
Using the raw data, the raw data after genome size adjustment, after PIA at genus level only, and after PIA at genus and higher taxa.

Species	Raw	GSA	PIA genus only	PIA genus + higher taxa
B. barbastellus	1.167315175	65.03001442	7.553769547	83.94423695
E. serotinus	17.44923858	4.42877912	38.69887744	91.86230937
M. alcathoe	98.00267797	13.22987095	95.65217391	98.48484848
M. bechsteinii	97.65533411	23.26961724	99.66058656	99.65753425
M. brandtii	96.41204582	6.666566394	89.44403545	97.8237882
M. daubentonii	95.22559083	4.449750271	100	100
M. mystacinus	97.03363276	15.02988482	100	99.75786925
M. nattereri	95.17923014	3.747882083	86.34128755	95.45731938
N. leisleri	0.446737775	8.795097698	24.81273483	73.84658685
N. noctula	1.253903435	26.61992359	10.02385587	77.3064054
P. auritus	4.443882283	86.40098459	45.19889634	91.56200975
P. austriacus	5.862758161	82.17403365	22.16477485	76.4211852
P. nathusii	6.813138038	40.81191139	36.03104644	86.77173053
P. pipistrellus	24.59797308	64.18058764	60.4539416	93.32369883
P. pygmaeus	16.19372605	62.09155843	10.69958358	81.32265132
R. ferrumequinum	46.25118361	8.797433932	92.18009547	97.4616813
R. hipposideros	94.50901437	85.17725172	83.25673552	92.71978022
Average	**46.97043425**	**35.34712635**	**58.95131735**	**90.45433149**

M. daubentonii had the highest success rate of sequence assignments after PIA (100%), whereas *N. leisleri* had the lowest (73.84%). This is likely due to the differences in database representation- Myotis is well represented, and Nyctalus poorly represented in the databases (see appendix E.3.2).

The Rhinolophid bats *R. hipposideros* and *R. Ferrumequinum* had 37.07 and 24.21 respectively assigned to Microchiroptera and Vespertilionidae

which is due to their previous grouping within these taxa (see section 1.6.2). After PIA, there were no Chiroptera reads in the blank library.

3.4.6.3 Chordata

5.79% of the sequences were returned as Chordata, not including the data reported as Chiroptera. The majority of these were mapped to *Homo sapiens*: which is hugely over-represented on the NCBI nt database, but may also be as a result of contamination of the samples, either during sampling (which was mostly done by ecologists without molecular biology training), or from the reagents or lab staff; there are 755 reads mapping to primates in the blank datasets (22.85% of blank reads). Apart from in *M. brandtii*, the number of reads mapping to primate in the bat datasets, is higher than in the blank libraries, suggesting that contamination in these samples originated in the sampling phases rather than in the lab, although this could also be an artefact of the carrier effect (Cooper, 1992). Future work should include a positive negative control, whereby DNA from another organism (not bat, arthropod, or human) is sequenced in parallel in order to fully quantify the impact of the carrier effect.

The second highest represented taxon was *Mus musculus*, which was not seen in either of the blank libraries, suggesting that these reads may be either misassignments (*M. musculus* is heavily represented in the Genbank nt databases), or that the guano samples came into contact with mouse DNA prior to sampling; many of the guano samples were collected from environments which would likely have been inhabited by mice.

The majority of the remaining Chordata reads are assigned higher up the tree, such as Eutheria.

Assignments of reads to non-chiropteran Chordata may be due to the database limitations. However, these sequences may also be as a result of environmental contamination, sampling contamination or library preparation. Some environmental and contamination is likely to be

inevitable: the only way to limit this would be to sample guano from captured bats in a sterile environment, which is beyond the scope of this project. Reagent or library preparation contamination should be limited to the taxa discussed in section 3.4.4.1 and would not explain the variability seen here.

3.4.6.4 Fungi

7.95% of the data returned were identified as from Fungi; however, none of these mapped to *Pseudogymnoascus destructans* (section 3.2.3) (Leopardi et al., 2015, Lorch et al., 2011, Puechmaille et al., 2011). The highest represented fungi species is *Penicillium rubens,* and is seen in all bat datasets (figure 3.17)*. P. rubens* is a model organism, and is thought to be the strain that Fleming observed to produce penicillin, rather than *P. chrysogenum*, as had been previously thought (Houbraken et al., 2011).

Histoplasma capsulatum, (Emmons et al., 1966, Miller, 1992) had only low numbers of reads (<250) found in a number of bat datasets. *Aspergillus* (Miller, 1992, Nováková, 2009) was found in all of the bat datasets, with high levels in four *M. mystacinus* samples: BatID 806, BatID 1325, BatID 2258, and BatID 2550, that came from across the range of *M. mystacinus* in Great Britain. The majority of bat species had Eurotiomycetes as the largest class of fungi, except for *M. alcathoe*, which had Microsporidia as the largest class, and *M. bechsteinii*, which had Saccharomycetes as the largest class.

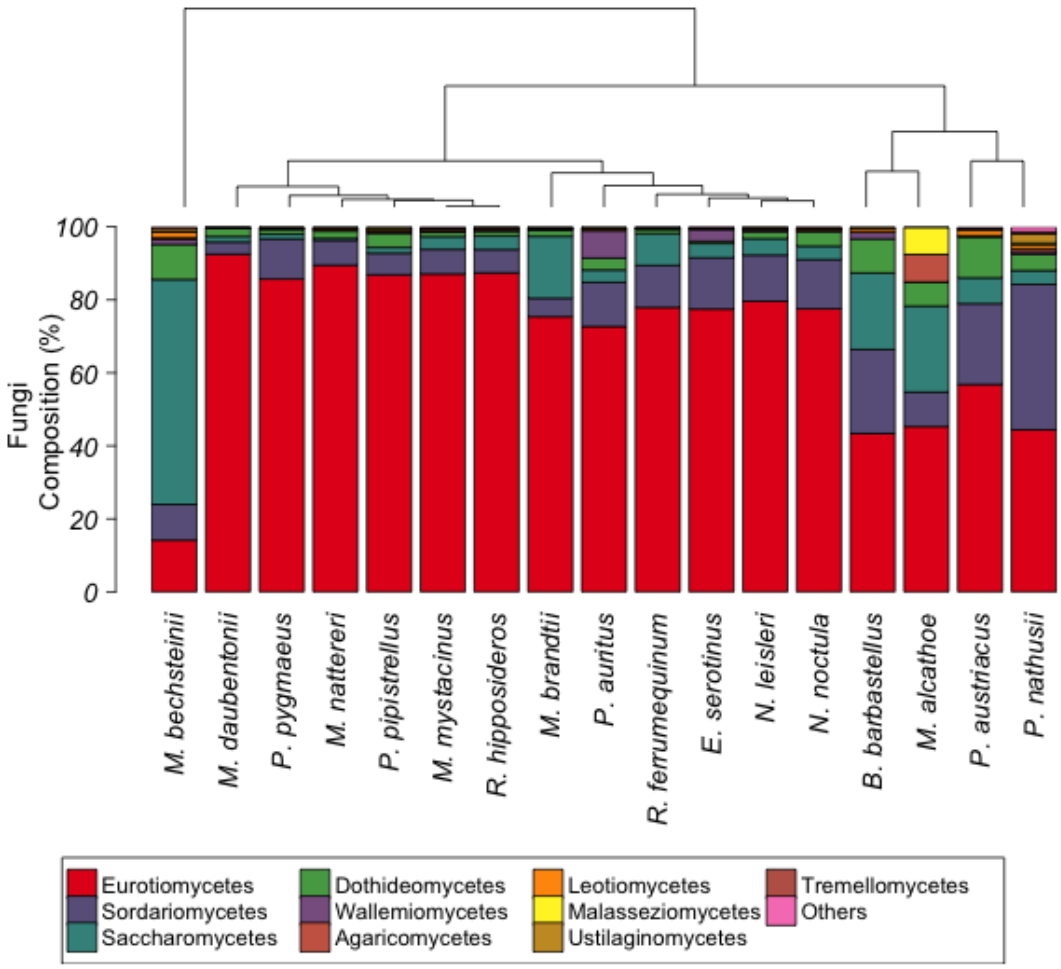


Figure 3.17: Fungi associated with each bat species

Data at assigned at class level. Data sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2.

Microsporidia have been identified as emerging pathogens, in particular the *Encephalitozoon* spp., which have been found to cause pathogenicity in bats (Childs-Sanford et al., 2006). High representation of Saccharomycetes in *M. bechsteinii* may be due to the combined effects of *S. cerevisiae* as a model organism, and the low number of fungi reads in *M. bechsteinii.*

As fungi cell walls do not easily chemically lyse, most DNA extraction procedures require ribolysis, and therefore fungal DNA will likely be under-represented in these data (Fredricks et al., 2005).

38,153 fungal reads were submitted for Phylogenetic intersection analysis (up to 500 reads per sample). 7,628 reads were assigned after analysis (19.99% retained). After PIA, the highest represented taxon was Aspergillus with 16.96% of the fungi reads. The previously highest represented class (Eurotiomycetes) was only identified to this level in 7.87% of reads. As a result of the high discard rate, fungal data may not be as robust as other types. The high discard rate may be due to the poor database coverage, or due to high genetic similarity between genera.

3.4.6.5 Viral

Due to the library preparation method, we mostly see dsDNA viruses (Baltimore classification group 1). Of these data, the majority of results were of phage origin, which have been previously seen to comprise a large fraction of the bat virome (Li et al., 2010).

A sizable portion (0.533%) of the data was assigned to viral origins, the majority of which (10,205 reads) came from one *P. pipistrellus* sample (BatID 2272) from the south west of England submitted to the Ecowarwicker Ecological Forensics service on the summer of 2013. Of the viral DNA, there are 9,286 of the filtered reads, which have been assigned to bacteriophages APSE 1-7, 8,708 of which came from the *P. pipistrellus* BatID 2272. APSE is a bacteriophage classified amongst the podoviridae (red, figure 3.18), or, in MEGAN, sometimes as 'unclassified phages'. APSE is an endosymbiont of *Hamiltonella defensa* a bacterium that is an endosymbiont of *Acyrthosiphon pisum*, the pea aphid (Degnan et al., 2009, Moran et al., 2005a, van der Wilk et al., 1999). There were 2230 sequences identified as *H. defensa* and 9 sequences identified as *A. pisum* in the BatID 2272 dataset. The symbionts APSE and *H. defensa* offer protection to the pea aphid against parasitic wasps (Moran et al., 2005a, Oliver et al., 2003). The increase in sequences that happens down the chain of the symbiotic relationships, suggests that there is a physical protective effect of the host on the endosymbiont DNA; the digestion of the *A. pisum* DNA is almost complete, but the *H. defensa* cells are partly protected from digestion by being encased in the *A.*

pisum, where they can reside within the haemocoel or in the bacteriocytes (Moran et al., 2005b, Sandström et al., 2001). A similar effect may be seen in the bacteriophage, which is may be protected in the cells of *H. defensa.*

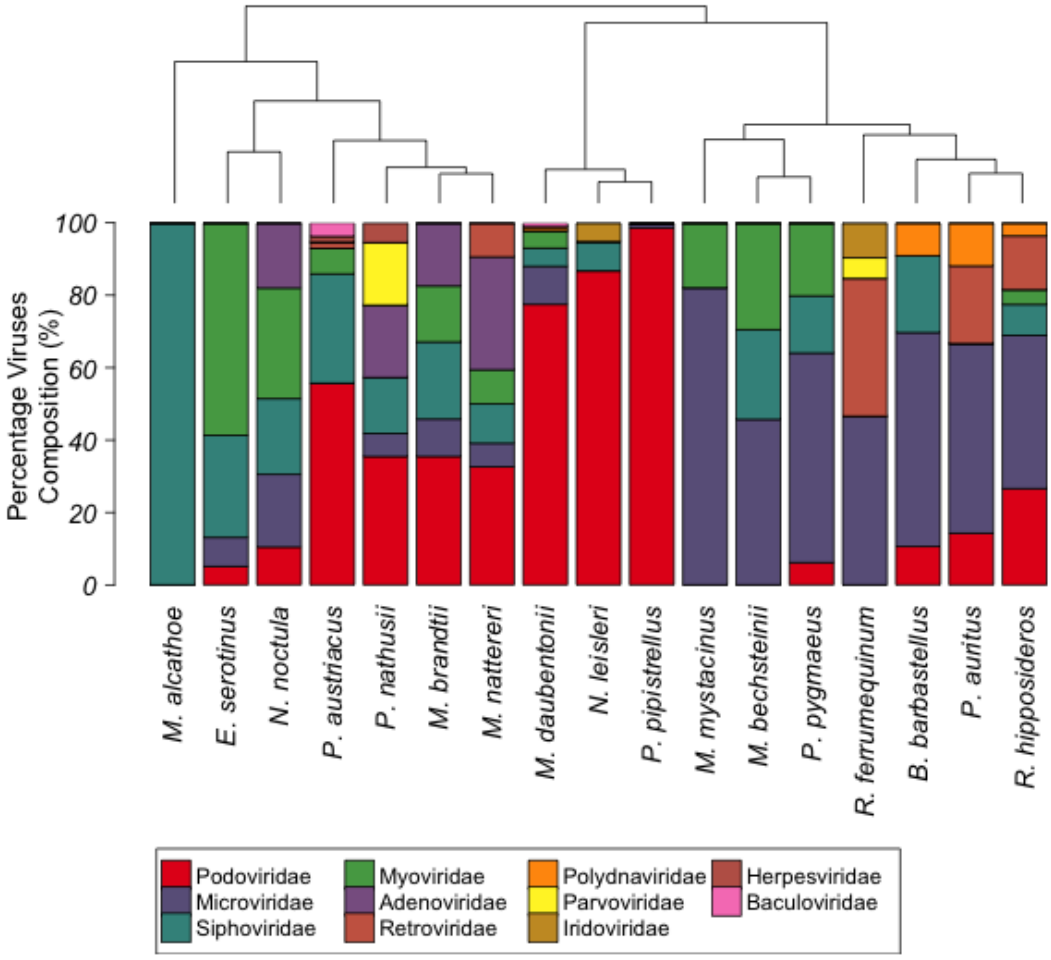


Figure 3.18: Viruses associated with each bat species

Data at assigned at "family" level. Data sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2.

Some ssDNA viruses were found in the datasets of many of the bats. Most of these map to Enterobacteria Phage PhiX174, and are seen in most of the bat datasets. This is most likely an artefact of the Illumina PhiX control that is used to aid cluster-formation during sequencing on the MiSeq. The rest map to Blattella germanica densovirus-like virus, which, whilst it is an ssDNA virus, requires a dsDNA intermediate during

replication. It was previously identified in Chinese insectivorous bats (Ge et al., 2012).

Of the pathogenic viruses discussed in section 3.2.3, the *Mastadenovirus,* and the *Gammaretrovirus,* are both represented in some of the datasets. The Mastadenovirus are in the Adenovirus family and have been identified in many vertebrates, including in several species of bats (Kohl et al., 2011). The gammaretroviruses are known to be involved in immunodeficiencies, leukaemias, and neurological diseases (Cui et al., 2012). Bat Gammaretrovirus DNA was found in low read numbers in the datasets of rhinolophid bats *R. ferrumequinum* (BatID 4232) and *R. hipposideros* (BatID 1134 and 4231).

Bats are known to carry a number of viral pathogens, however, none were found in these data. This is most likely due to the fact that many of the viral pathogens, (such as lyssaviruses, the causative agent of rabies) are RNA viruses.

20,059 viral reads were submitted for Phylogenetic Intersection Analysis (all viral reads in the dataset). Of these, 6,148 were assigned after analysis (30.64% retained). The majority of the reads (5,107) were assigned to “unclassified bacteriophages”. The lower retention of viral data after PIA may be due to the truncated nature of viral taxonomy (King, 2011). As a result of this, I suggest that PIA is not useful for use with viral data.

3.4.6.6. Other DNA sources

There are a number of other, smaller contributors to the total DNA sequences, including the Viridiplantae, which comprise 0.31% of the full dataset. The majority of the plant reads were assigned to the Poales, which may have originated from the environment from which the sample was collected, or from the diet of the arthropods eaten by the bats.

7,335 reads from the Viridiplantae were submitted for Phylogenetic Intersection Analysis (all Viridiplantae reads in the dataset). 1,229 were assigned after analysis (16.76% retained). A large proportion of the retained reads (435) were from a *M. daubentonii* sample (bat ID 1113). The high proportion retained in this sample, and the high proportion discarded in other samples suggests that these are true identifications, and highlights the power of PIA.

3.4.7. Dietary composition of guano

The major source for arthropod DNA from guano would most likely from dietary origins. 1.01% (26,438 reads) of the full dataset was assigned to Arthropoda. Reads were assigned to 19 orders of Arthropoda (there were 72 orders from Animalia, the highest represented of which was the Chiroptera), and to 1,114 Arthropoda species. Figure 3.19 shows the breakdown of the arthropod orders seen across the full dataset.

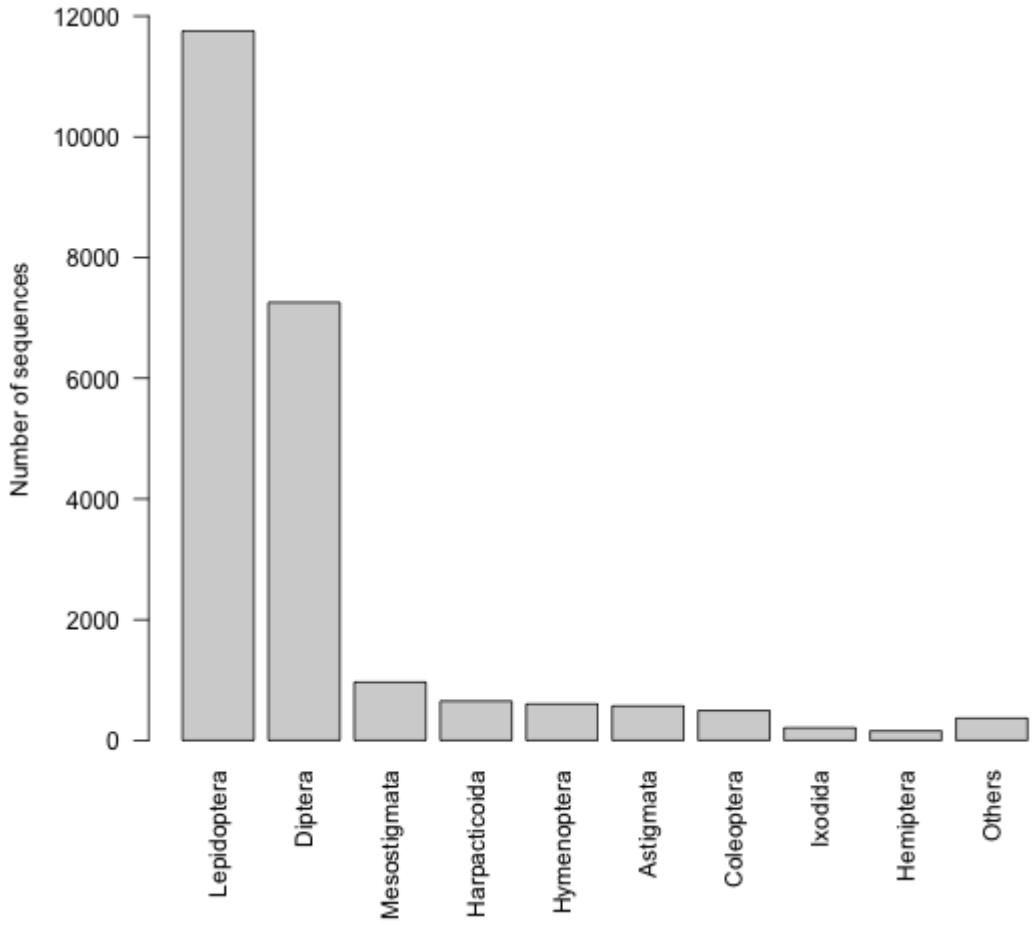


Figure 3.19: Assignment of arthropod sequences for all samples
These are shown to the level of order. Code in appendix D.2.10.

As was seen in previous studies (see chapters one and two), Lepidoptera and Diptera comprise the majority of bat diet species (figure 3.19). The largest order is the Lepidoptera, with 11,754 reads. The second largest order, the Diptera, has 7,256 reads.

3.4.8. Variation in diet between bat species

Figure 3.20 shows the variation in diets of all of the bat species collapsed to order level. The major distinguisher between dietary types is a preference for either Lepidoptera or Diptera. It is important to note that due to the limited number of samples used, each species diet will likely not be a true representation of the breadth of the diets of each species. This is further discussed in section 7.1.6.

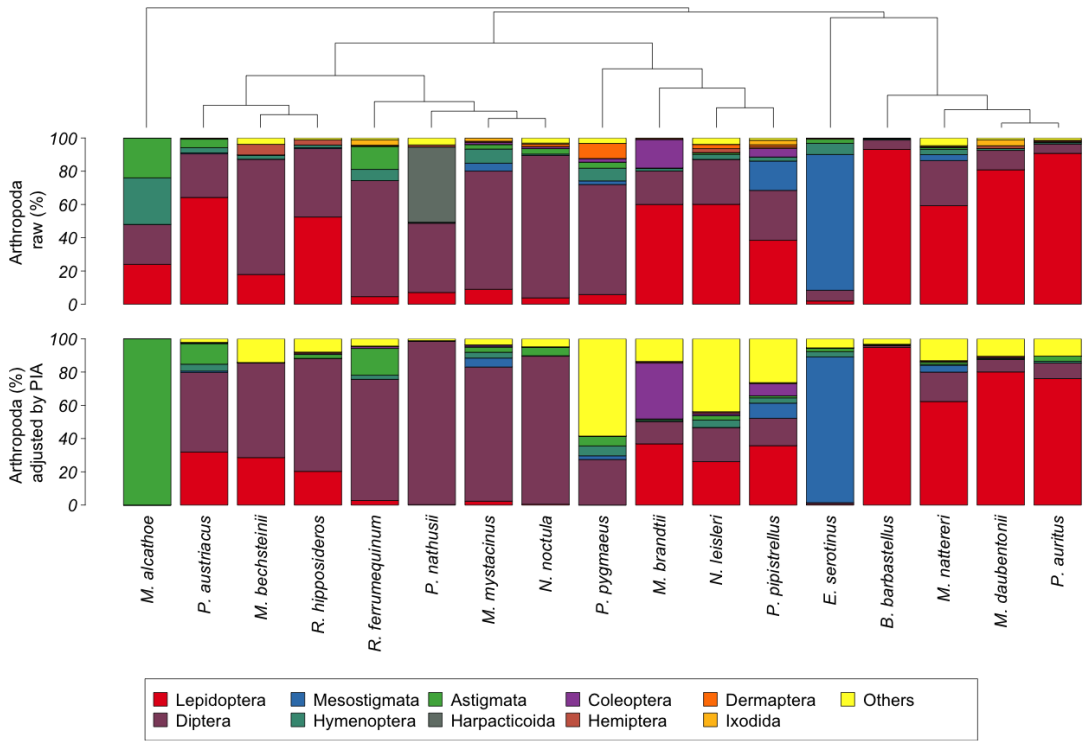


Figure 3.20: The diet of each bat species at order level before (a.) and after (b.) PIA
The average diets of each of the species. Prey taxa have been grouped by order (or higher level where assigned after PIA). “Others” includes Arachnida, Aranae, Blattaria, Cryptostigmats, Decapoda, Ephemeroptera, Insecta, Isopoda, Mantodea, Neuroptera, Opiliones, Pantopoda, Plecoptera, Psocoptera, Raphidiodea, Scorpiones, Siphonaptera, and Trichoptera. Diets sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2.

PIA analysis of the Arthropoda data largely supports the breakdown of the bat diets (figure 3.20), although it resulted in the discarding of 70.4% of the Arthropoda data (table 3.11). The phylogenetic intersection analysis has an estimated accuracy of 81% when assigning data from a shotgun metagenome (Smith et al., 2015). In section 3.4.6.2 (table 3.12) we found that in this study, the PIA had an estimated accuracy of 90.45%. These data will be compared to the metabarcoding dataset in chapter five.

Appendix C.3.4 contains a figure showing the diets of each bat species assigned at species level. *E. serotinus* was distinguished by its high

proportion of Mesostigmata reads (582), most of which came from BatID sample 2297.

Table 3.12: Data assignment before and after phylogenetic intersection analysis

Species	Number of Arthropoda BLAST hits before PIA	Number of Arthropod BLAST hits after PIA
B. barbastellus	3597	708
Blank	21	5
E. serotinus	824	625
M. alcathoe	36	4
M. bechsteinii	355	14
M. brandtii	2301	639
M. daubentonii	1879	311
M. mystacinus	981	554
M. nattereri	1624	520
N. leisleri	680	183
N. noctula	1640	787
P. auritus	1295	101
P. austriacus	4001	1220
P. nathusii	1577	478
P. pipistrellus	2314	508
P. pygmaeus	498	137
R. cont	23	1
R. ferrumequinum	1921	880
R. hipposideros	1015	191

4.99% of the Arthropoda data was assigned to the mt protein coding region (Cameron and Whiting, 2008), which is equivalent to around 0.069% of the dataset being assigned to barcoding regions. This indicates that the genetic space that is exploitable by shotgun metagenome approaches, and thus the sensitivity, is high.

3.4.9. Genome size adjustment

Figure 3.21 (a.) shows the proportion of data assigned at order level before and (b.) after genome size adjustment.

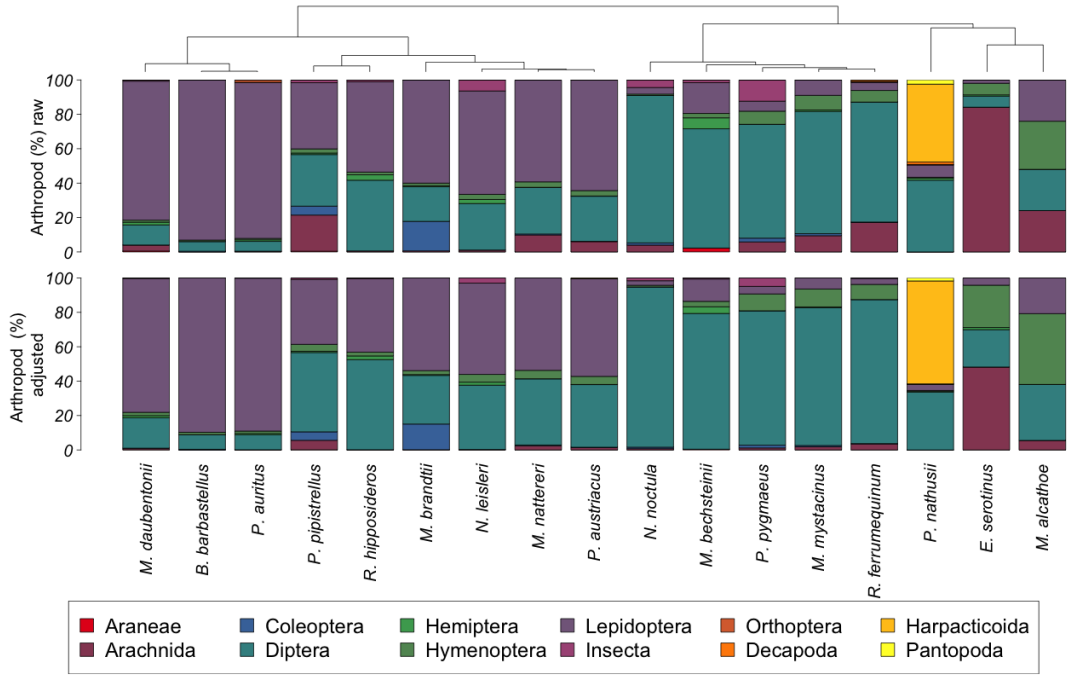


Figure 3.21: Data assignment after genome size analysis

a. Before adjustment, b. after adjustment. Code adapted from D.2.2.

Most diets do not change dramatically after genome size adjustment, except for that of *E. serotinus,* which had a high proportion of the DNA sequences in its dataset mapping to Arachnida; specifically to the Mesostigmata (see section 3.4.8), which were not represented in the genome size spread-sheet (no genome size information was available at the time of study; 2015), as a result, the reads were adjusted by the average genome size of the Arachnida, which have, on average, larger genome sizes than the other arthropods (Gregory and Shorthouse, 2003). Consequently, whilst a high proportion of the *E. serotinus* reads are assigned as Arachnida, a smaller proportion of the diet may be from arachnids, as their large genome sizes can cause them to be over-represented. The Lepidoptera reads are reduced, due to their larger genomes, and the Diptera proportion is increased, due to their smaller genomes. We did not adjust by organism sizes as the overall biomass of an order in the diet is the most important factor, whether it is one or many organisms. Genome size adjustment was also carried out on the post PIA data (figure 3.22).

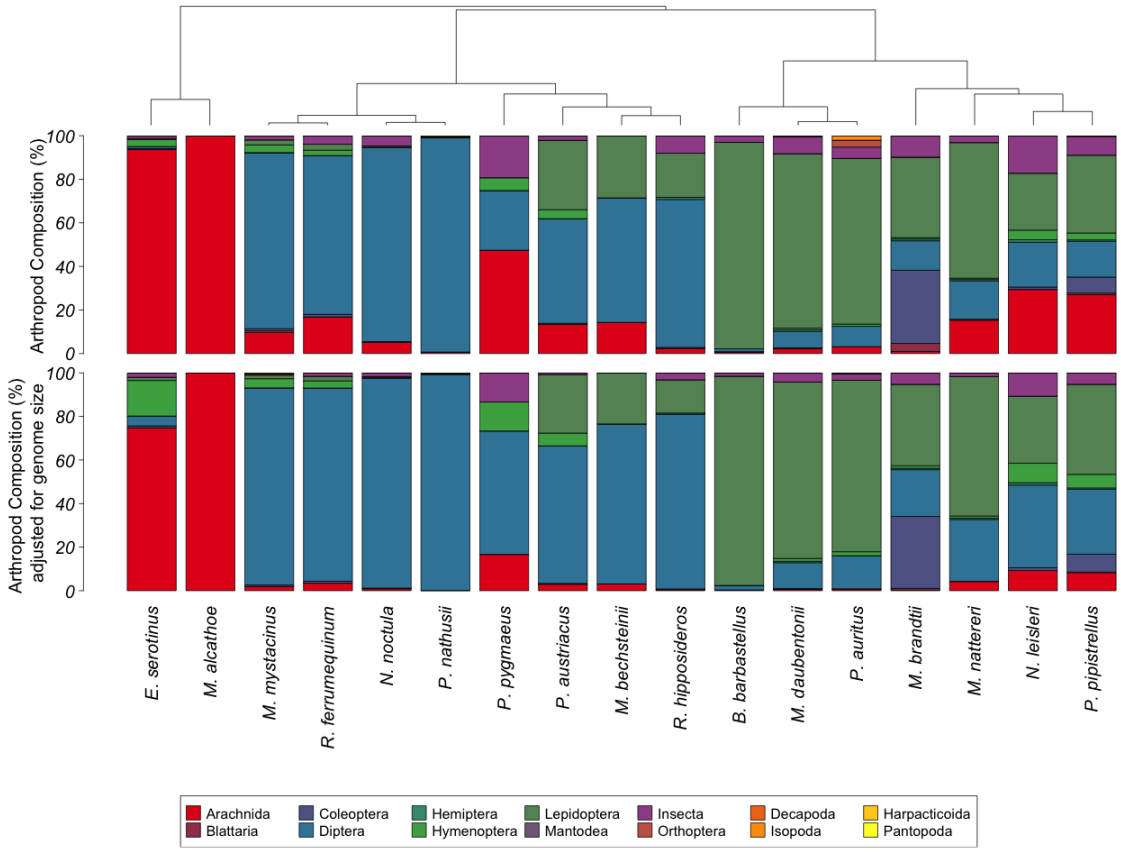


Figure 3.22: Data assignment after PIA before and after genome size analysis
a. Before adjustment, b. after adjustment. Code adapted from D.2.2.

As seen previously (figure 3.20), the major adjustment was to the Arachnida (red), which were enriched after GSA in all of the datasets they are found in. As a result of this analysis, we propose that the Arachnida biomass in the diet is somewhat greater than is suggested by the raw data. This may also be the case with the Hymenoptera (figures 3.20 and 3.21).

3.4.10. Dietary diversity, niche breadth, and overlap

Figure 3.23 shows the dietary diversity (Shannon-Weaver's diversity index, grey bars), and the niche breadth (Levin's index, black bars). These were calculated using all data assigned at order level (discounting anything assigned only to class or phylum, and collapsing to order anything assigned at family, genus, or species). The unadjusted data

(without PIA or genome size adjustment) was used as both of these adjustments had broadly supported the unadjusted data.

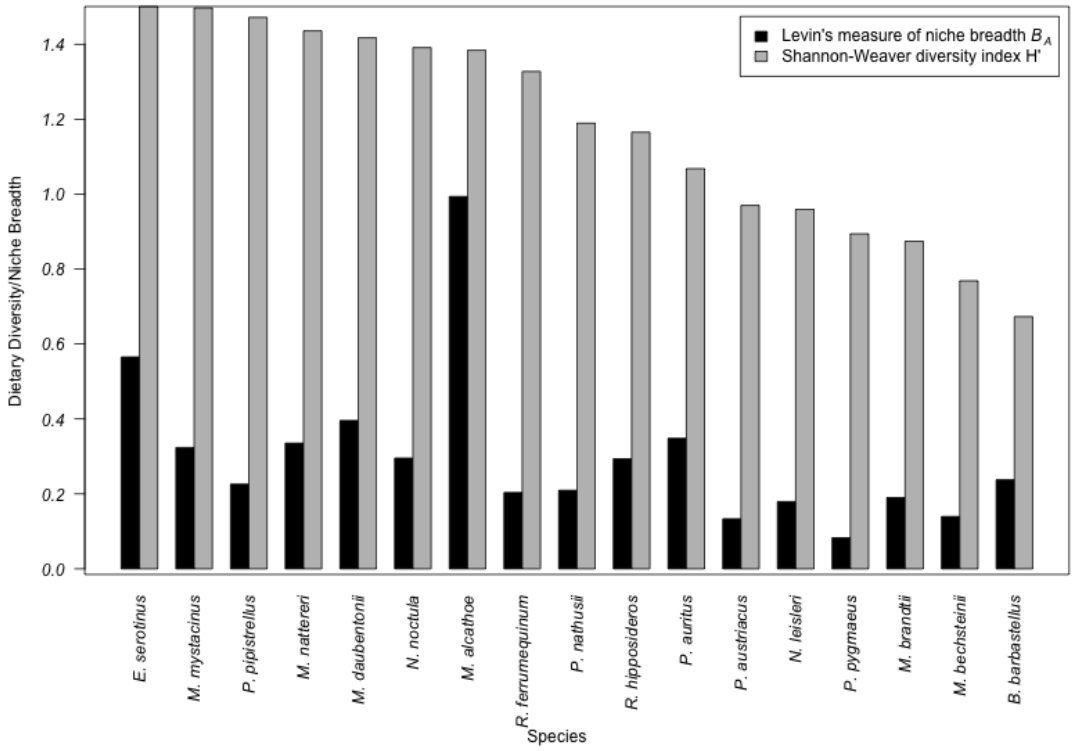


Figure 3.23: Dietary diversity and niche breadth of each bat species

The dietary diversity calculated using Shannon-Weaver diversity index (H) (grey) and niche breadth calculated using Levins standardised index (B_A) (black). Arthropod data assigned at order level. Code adapted from appendix D.2.3, values in appendix E.3.4.

There is wide variety in the diversity and niche breadth of the Great British bat species (figure 3.23). *B. barbastellus* with a Shannon-Weaver diversity index of 0.67, has the smallest diversity, whereas *E. serotinus* with an index of 1.5, has the greatest diversity. The species with the smallest niche is *P. pygmaeus,* with a standardised Levin's measure of 0.08, whereas the species with the broadest niche is also *M. alcathoe* with a measure of 0.99. *P. pygmaeus* and *P. austriacus* have the lowest dietary breadths, suggesting that they may be vulnerable to extinction pressures (Boyles and Storm, 2007, Safi and Kerth, 2004). Due to the small sample sizes, robust statistical tests are inconclusive, however, future work should develop this.

As was seen in the literature (chapter two), the Lepidoptera and Diptera form the basis of the diets of the Great British bats studied (figures 3.19 and 3.20). This allows the formation of two broad dietary guilds (figure 3.24). The first has higher proportions of Diptera and lower proportions of Lepidoptera. This guild contains all three of the *Pipistrellus* species, both of the *Rhinolophus* species, in addition to *M. bechsteinii, N. noctula, M. mystacinus* and *P. austriacus.* The second guild contains *N. leisleri, B. barbastellus, P. auritus, M. alcathoe, M. nattereri, M. brandtii,* and *M. daubentonii* all of which have a larger proportion of their diets comprised of the Lepidoptera. *E. serotinus* do not cluster into either of these guilds, as they feed heavily (>40%) on Mesostigmata.

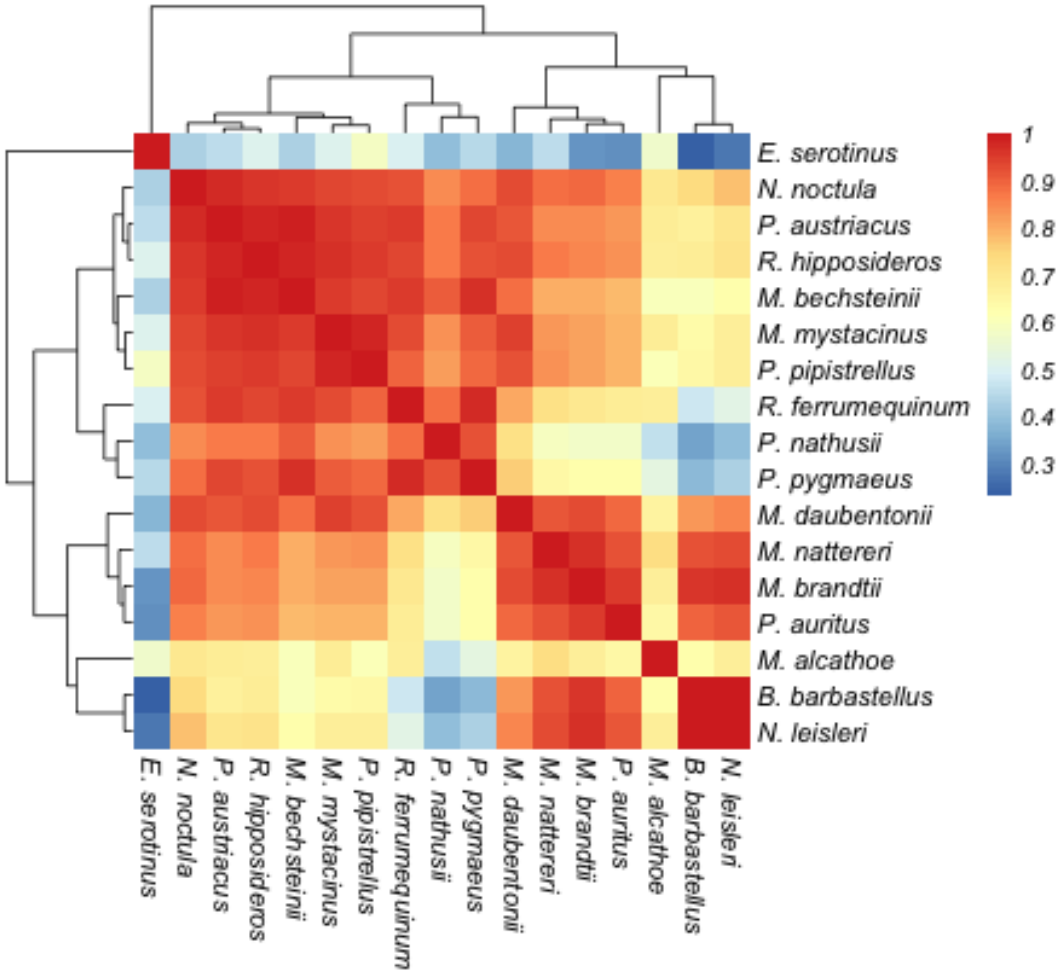


Figure 3.24: The dietary overlap of each bat species with Arthropoda at order level Calculated using Pianka's index of niche overlap. A value of 1 (red) suggests that the diets are identical, whilst a value of 0 (blue) indicates that there is no overlap. Species have been hierarchically clustered by niche overlap similarity. Code adapted from appendix D.2.4., values in appendix E.3.4.

The mean overlap was 0.768, with the highest overlap (0.99) being between *B. barbastellus* and *N. leisleri*, and the lowest overlap (0.23) between *E. serotinus* and *B. barbastellus*.

B. barbastellus and *N. leisleri* appear to be in direct competition (overlap of 0.994) for a small dietary range (dietary breadths of 0.237 and 0.179 respectively). This may go some way to explaining the decreasing population trend reported and near threatened categorisation by the IUCN for *B. barbastellus* (I.U.C.N., 2013), whereas *N. leisleri* is categorised as least concern. As a result, conservation efforts should focus on supporting *B. barbastellus* populations.

M. brandtii appears to be vulnerable, having low dietary diversity and niche breadth, in addition to being in direct competition with *M. daubentonii, M. nattereri,* and *P. auritus.* All of the competitors have more robust dietary diversities than *M. brandtii* (figure 3.23), and have stable (*M. nattereri* and *P. auritus*) or increasing (*M. daubentonii*) IUCN profiles (table 1.3) (I.U.C.N., 2013). *M. daubentonii* are thought to exploit different habitats; often feeding over water, however, *P. auritus*, due to their commonness, may be antagonistic to *M. brandtii*.

Additionally, the *P. pygmaeus* appear to compete with both the *R. ferrumequinum* and *P. nathusii*. Both the *R. ferrumequinum* and *P. nathusii* are considered to be rare in Great Britain, suggesting that *P. pygmaeus* have expanded into a niche that has not been exploited by more common bat species. Furthermore, it is possible that, due to the decreasing population trends in *R. ferrumequinum* (I.U.C.N., 2013), this niche had been previously occupied by *R. ferrumequinum*, but that in their absence, *P. pygmaeus* has expanded into this niche.

E. serotinus and *M. alcathoe* appear to have their own niches, suggesting that not all niches within Great Britain have been exploited by Great British bats.

3.5. Discussion

3.5.1. Guano quality

With age, exposure to the elements, and to attack by bacteria, fungi, and Arthropoda, the quality of the DNA in a sample would be reduced due to DNA fragmentation (3.3.1.4.1).

The Muscidae (Diptera) are seen in libraries from *M. brandtii, N. noctula, R. hipposideros,* and *R. ferrumequinum.* The Muscidae DNA may have originated from dietary sources as all of these bats were recorded as feeding on other dipteran species, however the Muscidae are often coprophageous (Waage, 1979), so these DNA may have come from arthropods feeding on the guano.

Other faecal decomposing arthropods include the Scatophagidae, Fanniidae, and Calliphoridae (Galante and Marcos-Garcia, 2004). Neither the Scatophagidae or the Fanniidae were seen in the full dataset, however, the Calliphoridae were represented in the datasets of some of the bats, and with reads seen in both *N. noctula* and *R. ferrumequinum.* Most of the Calliphoridae and Muscidae reads from *N. noctula* came from one sample, BatID 1839, and in *R. ferrumequinum* from BatID 1111.

Despite the presence of coprophageous arthropods and environmental fungi, which may be predictors of poor guano quality, the long read length of all of the sequence data (figure 3.8) implies that the DNA was of sufficient quality for this study.

A large proportion of the DNA sequenced was found to be of bacterial origin (see section 3.4.6.1). Bacteria may have originated from guts of the bat (including arthropod associated bacteria), from the environment that the guano was deposited, or from contamination. In this study we have focused on bacterial phyla previously identified as being present within bat guano (De Mandal et al., 2015), in order to identify bacteria that were of gut origins (rather than from the environment). We can also exclude bacterial DNA from contamination by the use of the blank library. Future

work should examine the remaining fraction of bacterial DNA as it may provide further information about the age of the samples or the environment that they were deposited in.

3.5.2. The impacts of hibernation on bat guano

Heterothermy is used by many mammalian species and some birds, in order to preserve energy in times of low food availability, or low temperatures. It involves the lowering of body temperature (T_b) and metabolic rate, and can occur in two forms; daily torpor, whereby short periods of heterothermy occur in each circadian cycle, and hibernation, whereby heterothermy continues for a number of days (Geiser and Ruf, 1995, Willis and Brigham, 2003). In preparation for hibernation, bats accumulate large reserves of fat in summer and early autumn (Speakman and Rowland, 1999). During periods of torpor, the body temperature fall below the optimum temperatures for the gut microflora, however regular periods of arousal occur even throughout cold periods (Speakman and Rowland, 1999).

Prior to hibernation, bats, and other mammals, will accumulate fat reserves by increasing the proportion of fat rich prey in their diets during the end of summer and autumn (Clare et al., 2014, Levin et al., 2013). This will be the main source of energy for the period of hibernation (Young, 1976). In this dataset, whilst not significantly different, the proportion of reads assigned to Arthropoda was highest in the samples collected in autumn (figure 3.25). The seasonal variation within the bat diets will be further discussed in chapters four, five and seven.

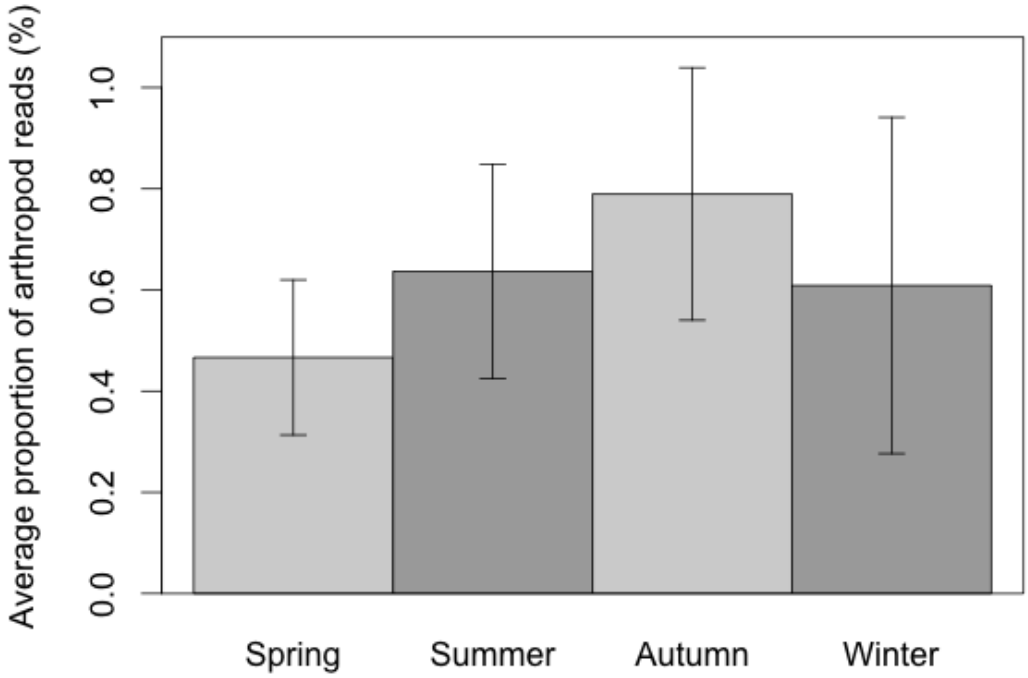


Figure 3.25: The proportions of Arthropoda reads over the seasons
Proportions of Arthropoda reads grouped into spring, summer, autumn and winter. Data are mean ±. SEM. Code adapted from appendix D.2.3. Numbers of samples: Spring- n=42, Summer- n=46, Autumn- n=51, Winter- n=30.

3.5.3.1. Modulation of the gut microbiota across the seasons

The structure of the microflora is strongly linked to the host diets; diets are key to supplying substrates for microbial growth, although host-derived nutrients are also used (Bäckhed et al., 2005, Carey et al., 2013). In return, the host receives nutrients, which would be otherwise inaccessible (Bäckhed et al., 2005, Dale and Moran, 2006).

The ratio between the Firmicutes and the Bacteroidetes is associated, in humans and other mammals, with energy absorption and obesity (Clemente et al., 2012, Sommer et al., 2016). The ratio of Firmicutes to Bacteroidetes is thought to also be reflected in hibernation patterns in bats (Carey et al., 2013, Sommer et al., 2016) and other mammals (Sommer et al., 2016). The Firmicutes play a key role in the degradation of polysaccharides (Flint et al., 2012), whereas the Bacteroidetes are generalists, capable of degrading a variety of carbohydrate sources (Martens et al., 2008, Salyers et al., 1977, Sonnenburg et al., 2005).

In the data collected in this study, the proportions of Firmicutes were lower in the winter and spring period, and the proportions of Bacteroidetes were lowest in summer (figure 3.26).

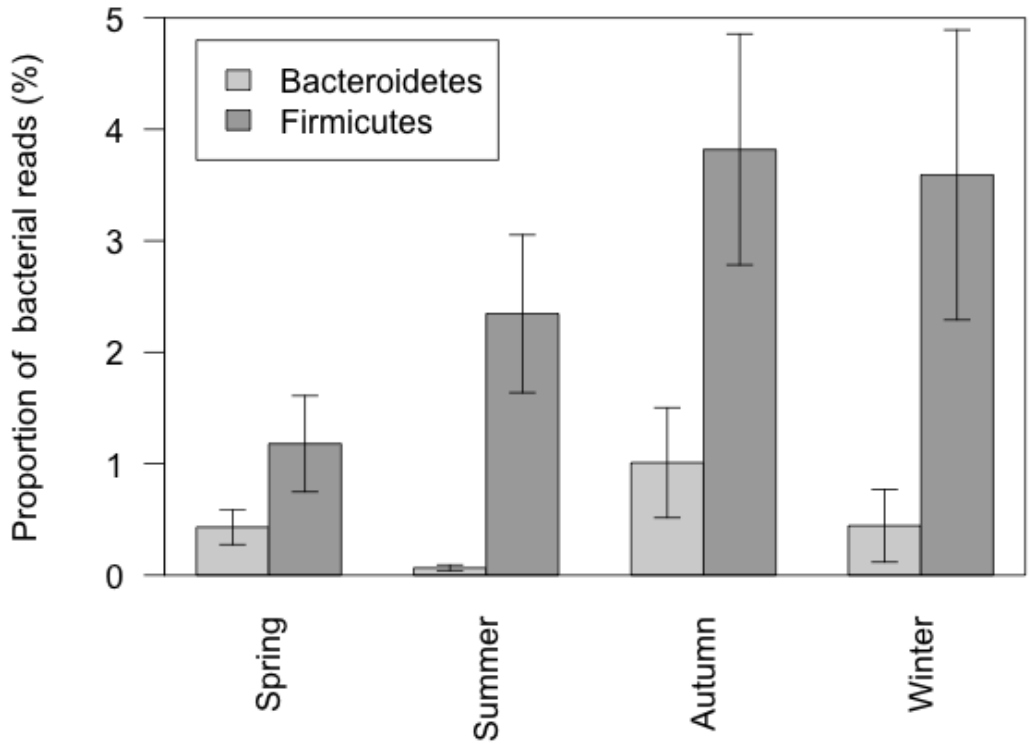


Figure 3.26: The proportions of Firmicutes and Bacteroidetes over the seasons
Proportions of bacteria grouped into spring, summer, autumn and winter. Data are mean ±. SEM. Code adapted from appendix D.2.3. Numbers of samples: Spring- n=42, Summer- n=46, Autumn- n=51, Winter- n=30.

In summer and autumn, polysaccharides from arthropod prey are abundant, and thus the proportions of Firmicutes increase, outcompeting the Bacteroidetes, which feed mainly on host-derived substrates. However, during hibernation, the availability of arthropod polysaccharides is limited, causing a decline in the Firmicutes, allowing for an increase in the Bacteroidetes which is still seen in spring.

A number of chitinase producing bacteria were seen in the data. The changes in proportions of chitinase producing bacteria through the

seasons can be seen in figure 3.27. The highest proportions of chitinase producing bacteria were found in guano samples collected in autumn, which may reflect the increased consumption of arthropods in preparation for hibernation (Whitaker et al., 2004). We then see a drop in chitinase producing bacteria after hibernation (after a slight lag). This trend is reflected in the proportions of Arthropoda reads in the dataset (figure 3.24). As in section 3.5.2, these data are not significantly different, however, future work should attempt to quantify this.

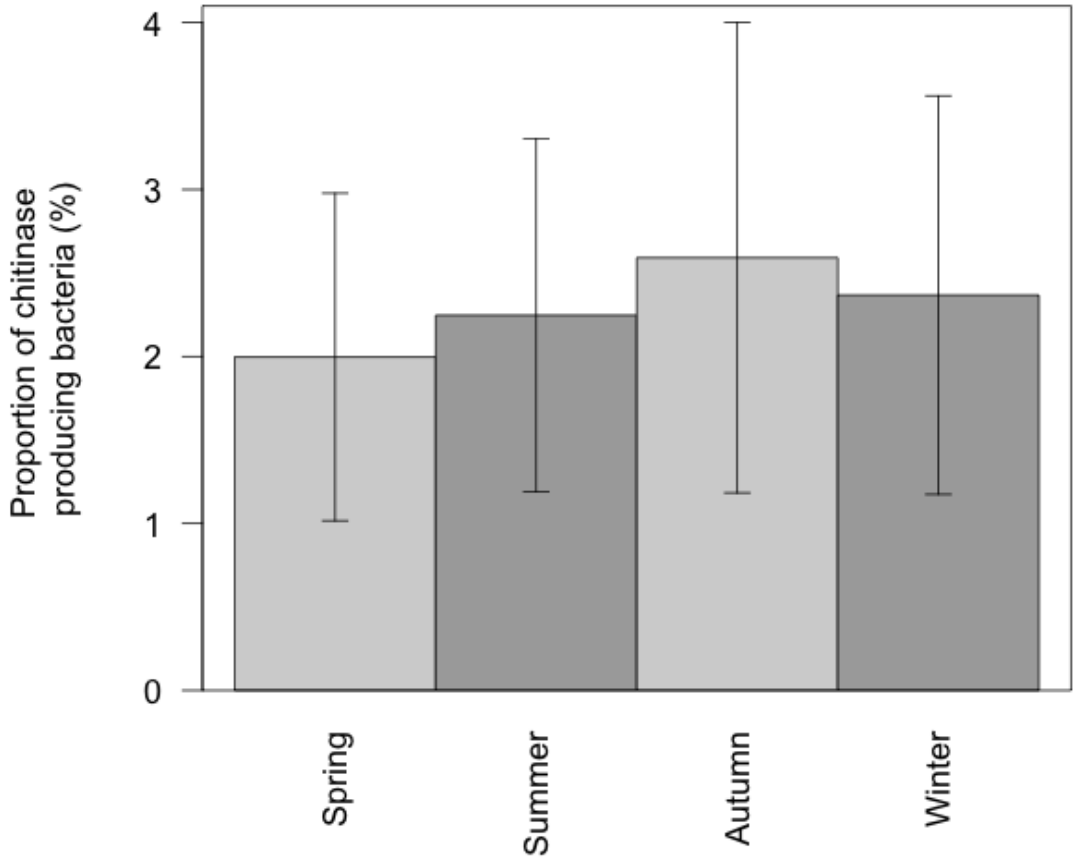


Figure 3.27: The proportions of chitinase producing bacteria over the seasons
Proportions of bacteria grouped into spring, summer, autumn and winter. Data taken from Aeromonas, Arthrobacter, Bacillus, Chromobacterium, Clostridia, Cytophaga, Enterobacter, Flavobacterium, Pseudomonas, Serratia, Streptomyces, and Vibrio. Data are mean ±. SEM. Code adapted from appendix D.2.3. Numbers of samples: Spring- n=42, Summer- n=46, Autumn- n=51, Winter- n=30.

C. perfringens produces chitinases (Lepp et al., 2010), but has also been identified as the primary cause of haemorrhagic diarrhoea in vespertilionid bats in Europe (Hajkova and Pikula, 2007, Mühldorfer et al.,

2011b, Mühldorfer, 2013). Enterobacter isolated from bats have been seen to have antibiotic resistance genes (Sherley et al., 2000). *P. putida* is a saprophytic soil bacterium (Nelson et al., 2002). This may be as a result of slower digestive transit during hibernation, allowing chitinase producing bacteria to access arthropod chitins.

3.5.3.2. Endothermy and fungal infections

Whilst we may know when the samples were collected, we rarely have information about when the samples were deposited. As a result, it is difficult to distinguish between the fungi which has originated from the bat gut and the fungi that has grown after deposition. We assume that pathogenic fungi are more likely to have originated from the bat, However, there was no evidence of known pathogenic fungi in any of the samples. It is possible that the peak seen in summer (figure 3.28) is as a result of warmer temperatures which results in the growth of fungi after deposition, and thus these fungi may have environmental origins.

Endothermy is thought to be a mechanism of inhibiting infection by fungal species, in particular psychrophilic (cold loving) fungi (Casadevall, 2012). Fungi have an optimal growth temperature of between 20 and 25°C, with little impact seen from temperature fluctuation (Dix, 2012). This is widely seen in the infection of *M. lucifugus* by *P. destructans* during their hibernation (Verant et al., 2012). This was reflected in this dataset (figure 3.28), where the proportion of reads assigned to fungi peaked in winter. We suggest that these fungi may be more likely to be of bat origin than the reads seen in summer.

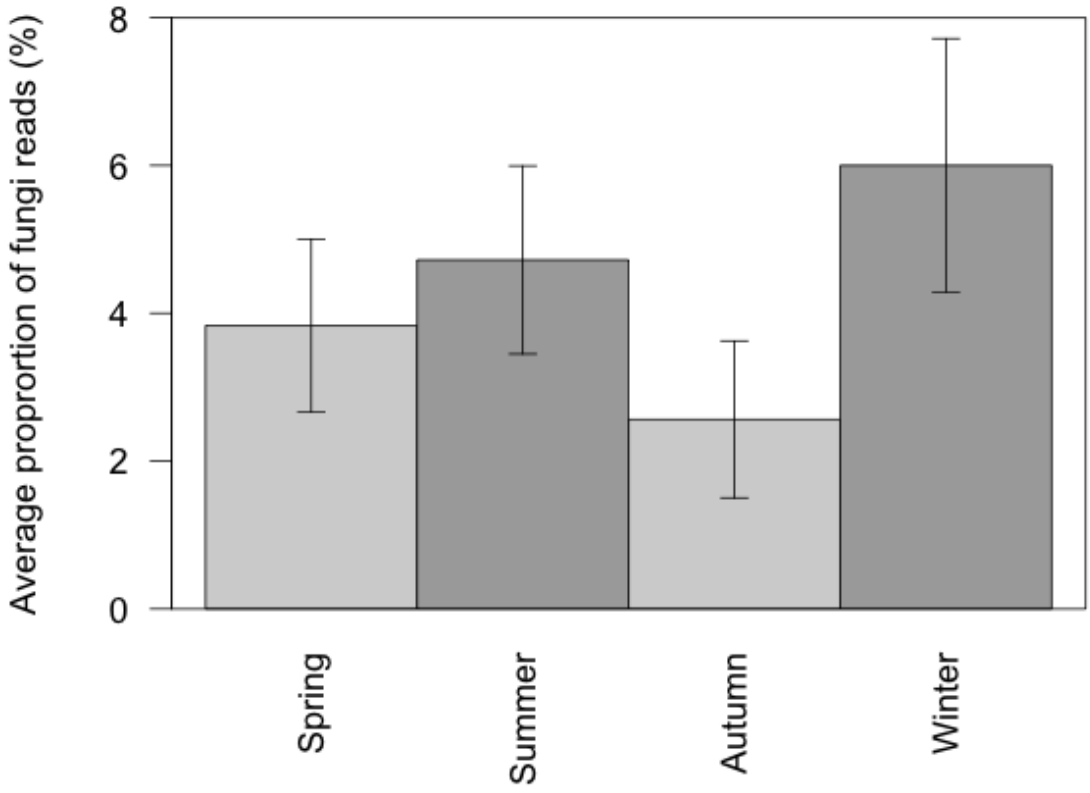


Figure 3.28: The proportions of fungi reads over the seasons

Proportions of Fungi reads grouped into spring, summer, autumn and winter. Data are mean ±. SEM. Code adapted from appendix D.2.3. Numbers of samples: Spring- n=42, Summer- n=46, Autumn- n=51, Winter- n=30.

3.5.4. Dietary overlaps and feeding styles

The high similarity between the diets of *N. leisleri* and *B. barbastellus* (Pianka's overlap of 0.99), may be due to the similarity of emergence times (18 and 19.5 minutes after sunset respectively) (Jones and Rydell, 1994, Russo et al., 2007). Additionally, both of these bat species are reported to be aerial hawkers (Holderied and Von Helversen, 2003), which is reflected in their diets, which comprise primarily of flighted arthropods; the majority of prey are Diptera, Lepidoptera. *N. leisleri* does have a small number of assigned to Astigmata, which is a flightless mite, however this is more likely to indicate grooming than gleaning. The high similarity of *N. leisleri* and *B. barbastellus* diets, suggests that these species, where they occur within the same habitat, may be in direct competition.

A number of the bat diets show some Arachnida (figure 3.21a); In particular *E. serotinus* has large proportions of Mesostigmata and Astigmata (figure 3.20) in its diets. The presence of mites in the diet is often thought to indicate grooming, rather than gleaning. The large genome sizes of Arachnids (Gregory and Shorthouse, 2003) is likely to artificially inflate the importance of mites in the diets, and the proportions of mites in the diets is greatly reduced after genome size adjustment (figure 3.21b) (see genome size spread-sheet in appendix E.3.5).

3.5.5. Guild structure and niche partitioning of cryptic species

Using the all of the data the overlap (Pianka's overlap index) between *P. pipistrellus* and *P. pygmaeus* was 0.887 (see appendix E.3.4.). This suggests that the diets of *P. pipistrellus* and *P. pygmaeus* may be less similar than previously thought; in the literature review (chapter 2) the Pianka's overlap was 0.93. In the case of the cryptic myotis species (*M. alcathoe, M. brandtii, M. mystacinus*), the average Pianka's overlaps between these species was 0.723. Whilst this is still considered to be "biologically significant" (Bethea et al., 2006), it is lower than that seen in previous studies (section 2.4.3).

3.5.6. Dietary breadth and extinction risk

In terms of dietary specialisation, it would be expected that the *E. serotinus*, being the greatest generalist, would be the most robust against variations in prey availability due to it's diverse diet and therefore be more successful than other species (Boyles and Storm, 2007). Additionally, E. serotinus appears to have very low competition from other bat species. This is reflected in it's "least concern" status on the IUCN red list, with it's stable population (see chapter 1, table 1.4) (I.U.C.N., 2013). Three of the four Great British bats with populations listed as declining (*B. barbastellus, M. bechsteinii,* and *R. ferrumequinum*) were among the four species with the lowest niche breadth. *P. pygmaeus* has the lowest niche breadth (0.08) and appears to be in direct completion to a number of other species (section 3.4.10), but its status is listed as unknown. This highlights the need for more information about the population trends of

this species, as this research suggests that it may be vulnerable to extinction pressures, and thus it may be prudent to consider amending its status as “least concern” (I.U.C.N., 2013). *M. brandtii’s* status should also be further investigated, as it also appears to be highly vulnerable (section 3.4.10).

The only Great British bat with a reported increase in population size is *M. daubentonii*. This is reflected in its relatively high niche breadth (0.396) and diversity (1.417).

Lepidoptera and Diptera are clearly vital in the diets of all bat species. This is reflected in figure 3.29, where Lepidoptera and Diptera explain most of the variation in principal components 1 and 2.

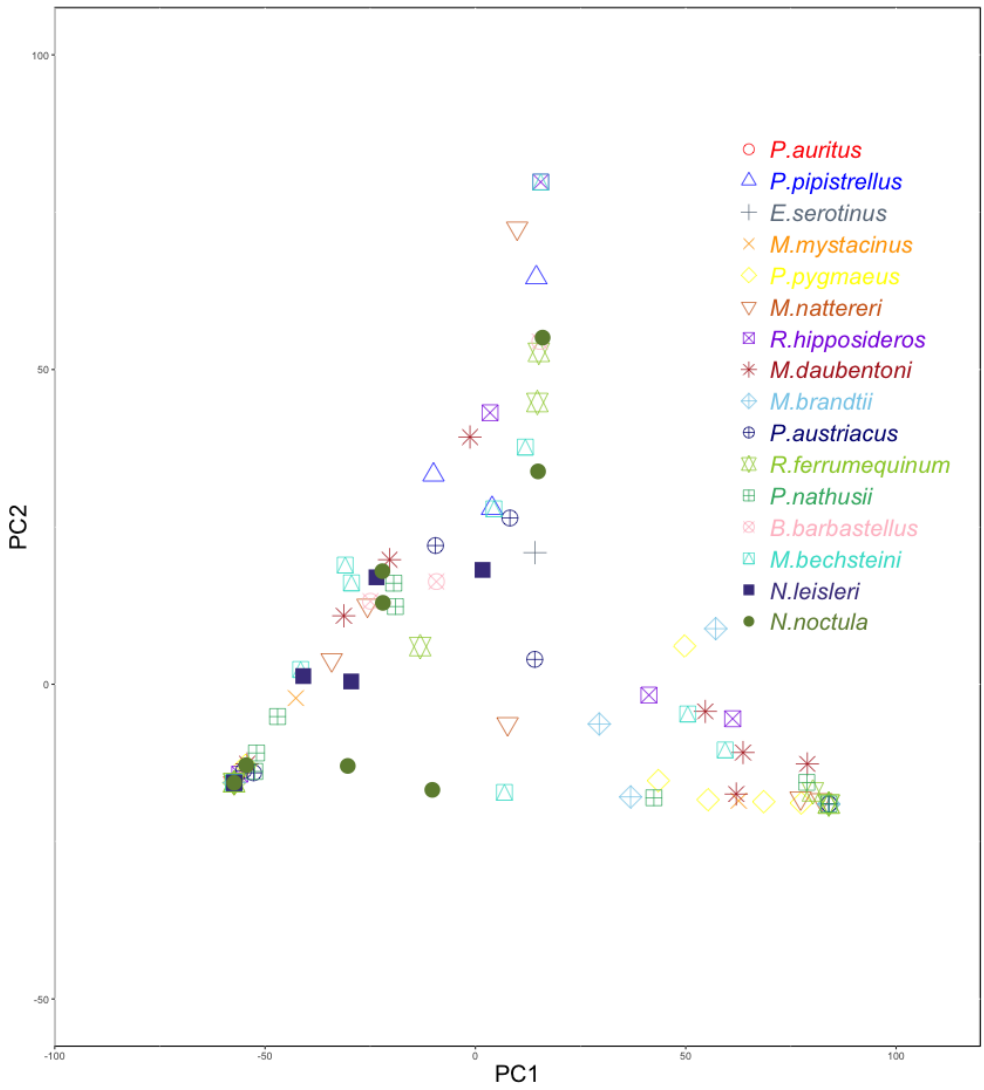


Figure 3.29: Distribution of the Great British bat diets in principal component space

Derived from the proportions of diet species within the libraries. Principal components one and two are shown; together they explain 74.45% of the variation (PC1- 50.31%, PC2- 24.14%). Component one is defined primarily by Lepidoptera (loading of 0.69) and Diptera (loading of -0.72), and component two by Coleoptera (0.1), Diptera (-0.6), Hymenoptera (0.14), Ixodida (0.35), Lepidoptera (-0.63), and Mesostigmata (0.22). A summary and loadings can be found in appendix E.3.4. Figure code adapted from D.2.11.

3.5.7. Future work

There are two major areas where there is scope for future work, based on the work in this chapter, which is not addressed in other sections of this thesis. Firstly, undertaking library preparations to enable the sequencing of other types of viruses (not just dsDNA), could give further insight into

the viral pathogens carried by bats. Secondly there is scope for further investigations into the impact of hibernation cycles on bacterial and fungal flora, perhaps using a targeted metabarcoding approach.

3.6. Conclusions

The hibernation cycle appears to be a key driver in the bacterial and fungal flora associated with bat guano. The increased consumption of arthropods in order to gain fat reserves for winter drives an increase in the proportion of chitin producing bacteria, and causes a skew in the ratio of Bacteroidetes and Firmicutes. During hibernation, the fall in body temperature causes the bats to have increased susceptibility to fungal infections. Bacterial infections seem to be limited to individual bats, and there is no indication form these data that there are wide spread infections of any bacterial species.

In chapter four the data from the first Illumina MiSeq will be used to direct the design of primers for metabarcoding. The Arthropoda data from this chapter will be compared to the data collected from the literature (chapter two) and from the metabarcoding (chapter four) in chapter five and chapter seven will place these data in an ecological context.

Chapter 4 : Metabarcoding data

4.1. Abstract

This chapter is a metabarcoding analysis of the Arthropoda DNA from a selection of guano samples from across the range of Great Britain's species. Metabarcoding is used to give a high-resolution picture of the Arthropods present within a bat's diet.

The Illumina MiSeq run returned 12.898 GB of FASTQ data before adapter removal, quality control, and primer trimming, and after, read one returned 2.618 GB of which 9,514,856 reads were assigned to Arthropoda (90.155%). 38 orders of Arthropoda were identified, of which, Lepidoptera and Diptera were the highest represented. The bat species with the highest dietary diversity and broadest niche breadth was *M. bechsteinii*, and the species with the lowest was *M. alcathoe.* These data resulted in the formation of three major dietary guilds. The proportions of Diptera in all bat diets were highest in winter and in Scotland.

These data are compared to those collected in the literature review (chapter two) and from Arthropoda data collected from the shotgun metagenome analysis (chapter three) in chapter five.

4.2. Introduction

Metabarcoding has been widely used to study mixed environmental communities. It is increasingly being used to study mammalian and avian diets by sequencing barcode regions of prey species. For land plant species, a combination of *rcbL* and *matK* are typically used (Giovino et al., 2014, Parmentier et al., 2013). For animal species in the diets, the cytochrome c oxidase 1 gene is commonly used (Ratnasingham and Hebert, 2007), and due to the comprehensive nature of the databases available, was selected for this study (section 4.3.1.1).

Metabarcoding has a number of advantages in comparison to shotgun metagenomics. The major advantage is that the proportion of data returned that originates from the species of interest (in this case Arthropoda), will be far higher as none of the non-target DNA will be sequences (such as the bacterial community) (Bon et al., 2012, Srivathsan et al., 2015). As a result, many samples can be multiplexed onto one sequencing run, minimising the experimental time and costs (De Barba et al., 2014). One of the disadvantages is that it uses less of the genetic material laid down after cell death, which is discussed in chapter 3 (section 3.3.2.5.2) (Smith et al., 2015).

4.2.1. Addressing the issues associated with metabarcoding

It is crucial that the primers used will amplify the target DNA only, and that all of the target species are amplified. This can pose a problem when designing the primers; some *a-priori* knowledge of the target organisms expected to be present in the sample is required. In this study, the data from the literature review, and from the first shotgun metagenome MiSeq run were used during the primer design process (see section 4.3.1.1). It was also important that, as far as possible, the primers would not amplify non-target species. From the metabarcoding data, a large proportion of the DNA present originated from the bats, so it was vital that the primers would not spuriously amplify the Chiroptera cytochrome C oxidase gene (or any other part of the Chiroptera genome). It was also important that primers were checked against databases of bacteria to ensure that there would be no amplification of bacterial DNA. In addition to optimising the specificity of the primers, it is also important to ensure that experimental conditions (thermal-cycling conditions, etc.) are optimised to prevent spurious amplifications. This is particularly important when using DNA from guano, which contains a broad variety of DNA sources.

Primers are designed based on knowledge of target sequences from sequence databases, and poorly studied taxa may be under represented in the databases. This means that these taxa can be missed during the primer design process, and as a result the primers may fail to amplify

DNA from these taxa despite their presence in the sample, leading to issues of bias. The data stored in sequences databases is expanding dramatically (Benson et al., 2012), meaning it is important to regularly check primers against sequence databases, to ensure both that the breadth of the target species are being amplified, and that the spurious amplifications are minimised.

The reference DNA sequence database (such as the NCBI's nt database (Altschul et al., 1990)) used to assign taxonomic classifications to the output sequences can have large impacts on the success of taxonomic classifications; using a patchy database in which many taxa are not represented will not only increase the chance of sequences being unassigned, but will increase the likelihood of sequences being misassigned (Smith et al., 2015). It is therefore important to select the database appropriately. In this study, we have used GenBank's NCBI nt database (Benson et al., 2000, Benson et al., 2012) and the Barcode of Life Database (Ratnasingham and Hebert, 2007).

These issues can be further confounded by stochastic errors introduced during the PCR enrichment, whereby some PCR templates may be preferentially amplified over others (Best et al., 2015). In order to overcome this, duplicate PCR reactions are undertaken for each sample, which are then pooled (Pearman et al., 2014).

4.2.2. Selection of sequencing technologies

For barcoding, read-length is not particularly important, provided that the sequencer is able to provide the full sequence of the barcode used; the most important factors are the accuracy of the sequencing and the amount of data created. Therefore, Illumina sequencing technologies were selected due to their high read number. MiSeq sequencing has longer read length chemistry than the HiSeq, allowing for the full barcode region to be fully sequenced. Furthermore, it is more cost effective in this context than the HiSeq, which has prohibitively expensive upfront cost (although it is cheaper per megabase returned). It also would likely result

in a large level of needless redundancy in the data. As in chapter three, multiple samples were multiplexed into one machine run.

4.2.3. Expected data yield

For a MiSeq V2 reagent chemistry run with a read length of 2x 250bp, the expected data yield is between 7.5 and 8.5 GB. Assuming that 7.5 GB is returned over 108 samples (including the negative control), ~ 69.4 MB of sequence data would be returned per sample. This is equivalent to approximately 367,900 reads per sample. This provides enough barcode sequence data to allow for detailed dietary and phylogenetic analysis.

4.3. Materials and methods

4.3.1. Wet Lab

4.3.1.1. Designing primers to target specific barcode regions

In this second stage of this project, primers were designed to provide taxonomic and phylogenetic information about the bats and their diets, informed by the data obtained from the first stage. Mitochondrial targets are used as they show enough variation to allow for species level identification and are also present in high copy numbers within a sample (Clare et al., 2009). The key issue with this type of approach is that it limits identifications to sequences on databases such as BOLD (the Barcode of Life Database) and NCBI; however, this can be avoided by assigning sequence data to OTUs (Operational Taxonomic Units) where species level (or higher taxonomic level) cannot be resolved.

Primers used to identify the bat species, and those for arthropod barcoding, ideally have short amplicons and amplify mitochondrial genes which are more stable than genes located within nuclear DNA, due to the high mitochondrial copy number, and also due to the mitochondria's double membrane. Furthermore, the PCR targets are well represented on the databases such as GenBank (Benson et al., 2000, Altschul et al., 1990, Ratnasingham and Hebert, 2007). Primers must be able to amplify the range of target species required. Where the target gene is not highly conserved across the target species, a mixture of degenerate primers

(i.e. some positions on the oligonucleotide may have one or more of the possible bases represented) may be used.

Cytochrome c oxidase subunit 1 (CO1) and cytochrome B (cytB) were selected as preliminary candidate gene targets for this study (Hebert et al., 2003, Ward et al., 2005). CO1 had 916,830 sequences for target species on the NCBI nt database as of August 2015, and 5,657,681 on the BOLD, whereas cytB has 213,491 on the NCBI nt database, and is not recorded on the BOLD (Ostell et al., 2004, Ratnasingham and Hebert, 2007), making CO1 the best choice for target gene.

To address the problems of DNA fragmentation, discussed in section 3.3.1.3.4, a short amplicon should be used (Golenberg et al., 1996). Previous barcoding studies of bat diets used Matthew Zeale's primers, ZBJ-ArtF1c and ZBJ-ArtR2c which amplify a 167bp region of the CO1 gene (n.b. not 157bp as is reported in the their paper (Zeale et al., 2011)). The forward primer (ZBJ-ArtF1c- 5'-AGATATTGGAACWTTATATTTTATTTTTGG-3') contains one degenerate base, W, which, when the primer is synthesised means that half of the primer oligonucleotides will have an A in that position and half T (Cornish-Bowden, 1985). As a result, only half the primer oligonucleotides would bind exactly to any one target COI. In the reverse primer (ZBJ-ArtR2c- 5'-WACTAATCAATTWCCAAATCCTCC-3'), there are two such degenerate positions (both W). As a result, the chance of a target molecule having an exact match to the target primer is around a quarter. Degeneracy at this scale is unlikely to be an issue, as some mismatching can be tolerated during PCR (which depends on the reaction conditions such as Mg2+ concentration, and the polymerase used (Markoulatos et al., 2002)). However, it means that other mismatches may cause the binding efficiency to drop, or to fail altogether (Sipos et al., 2007).

We undertook BLAST searches (Altschul et al., 1990) of both of the primers using the MegaBLAST algorithm (Zhang et al., 2000) on the

nucleotide database, to check the likelihood of amplification success. A list of potential prey taxa was compiled from the literature review and from the first MiSeq run. Query cover percent and percentage identity for each of the top hits was recorded. These were summed to give an exact match score of 200. Due to the degenerate bases, the highest possible score for the forward was 197, and the highest for the reverse was 192. 60.95% of the expected taxa had exact matches to ZBJ-ArtF1c, and 85.71% had exact matches to ZBJ-ArtR2c. As these were the top matches for each of the taxa, it is likely that the proportions of exact matches to the full dataset is lower than this. As a result, it is unlikely that these primers would be sufficient to amplify the full breadth of the Arthropoda, if they are present in the guano.

The sequences of the top taxa that had mismatches to the Zeale primers were downloaded. These were aligned using ClustalX (Thompson et al., 2002), and primers were manually designed to match the breadth of these sequences. There are a number of criteria for primer design: they should not have long repeats of any one base, they should not be palindromic, all of the primers should have a similar T_m (°C), they should not have too many (>2) degenerate bases, they should be 18-22bp long, and ideally should have a GC content of around 50% (calculated using an approximation described in eqn 1).

$$T_M = \big((A+T)\times 2\big) + \big((G+C)\times 4\big)$$

(Eqn 1, Melting temperature)

No primer had more than 4 fold degeneracy; either one degenerate base which can encode all four (N) or three bases (B, D, H, or V), or up to two degenerate bases which can each encode two bases (e.g. R, Y, S, G or T) (Cornish-Bowden, 1985, Markoulatos et al., 2002, Sipos et al., 2007). As a result, a system of 19 forward and 22 reverse primers was developed (table 4.2), which, between them are degenerate enough to amplify all of the target arthropods (and any in-between), but should not amplify non-arthropod COI targets. In order to constrain the melting

temperatures, the forward primers were slightly shorter than that used in Zeale's study (Zeale et al., 2011).

Table 4.1: Arthropoda COI primers

These amplify a 167bp region of the Arthropod COI gene, and are based on the Zeale primers ZBJ-ArtF1c and ZBJ-ArtR2c (Zeale et al., 2011).

Primer Name	Orientation	Sequence	Tm° C	CG%
>1_F_Arthropoda_COI	Forward	AGACATTGGAACWTTATATTTC	54.1	27.3
>2_F_Arthropoda_COI	Forward	AGATATYGGAACTTTATATTTT	51.7	20.5
>3_F_Arthropoda_COI	Forward	AGATATTGSTACTTTATATTTT	48.6	18.2
>4_F_Arthropoda_COI	Forward	AGATATTGGAACACTATATTTT	51.2	22.7
>5_F_Arthropoda_COI	Forward	AGATATTGGAACWTTATATTTT	50.9	18.2
>6_F_Arthropoda_COI	Forward	AGATATTGGAACMTTATATTWA	50.5	20.5
>7_F_Arthropoda_COI	Forward	AGATATTGGAACTWTWTATTTT	50.5	18.2
>8_F_Arthropoda_COI	Forward	AGATATTGGWACTTTATACTTT	49.5	22.7
>9_F_Arthropoda_COI	Forward	AGATATTGGAACTTTATACTTT	50.6	22.7
>10_F_Arthropoda_COI	Forward	AGATATTGGMACTTTATATTTT	52.1	20.5
>11_F_Arthropoda_COI	Forward	AGATATTGGAATWTTATATTTT	49.5	13.6
>12_F_Arthropoda_COI	Forward	AGATATTGGAAYATTATAYTTT	49.8	18.2
>13_F_Arthropoda_COI	Forward	AGATATTGGRACATTATATTTT	52.2	20.5
>14_F_Arthropoda_COI	Forward	ARATATTGGWACATTATATTTT	50.1	15.9
>15_F_Arthropoda_COI	Forward	GGATATCGGAACCCTATATTT	59.1	36.4
>16_F_Arthropoda_COI	Forward	GGATATTGGAACATTATATTTT	53.6	22.7
>17_F_Arthropoda_COI	Forward	GGATATTGGAACTTTRTATTTT	54.5	25
>18_F_Arthropoda_COI	Forward	RGATATTGGAACTATATATATA	45.8	20.5
>19_F_Arthropoda_COI	Forward	WGATATTGGAACWTTATATTTT	51.6	18.2
>1_R_Arthropoda_COI	Reverse	AACGAGTCAATTACCAAATCC	59.8	38.1
>2_R_Arthropoda_COI	Reverse	AACTAATCAATTYCCAAATCC	57.7	31
>3_R_Arthropoda_COI	Reverse	MACTAATCAATTTCCAAAYCC	58.4	33.3
>4_R_Arthropoda_COI	Reverse	WACTAATCAATTTCCAAATCC	56	28.6
>5_R_Arthropoda_COI	Reverse	AACTAATCAGTTWCCAAATCC	55.6	33.3
>6_R_Arthropoda_COI	Reverse	AAKAAGYCAATTTCCAAATCC	60	33.3
>7_R_Arthropoda_COI	Reverse	AATTAAYCAATTTCCAAAWCC	57.6	26.2
>8_R_Arthropoda_COI	Reverse	DATTAATCAATTACCAAATCC	54.1	25.4
>9_R_Arthropoda_COI	Reverse	AATTAGTCAATTTCCAAATCC	56.7	28.6
>10_R_Arthropoda_COI	Reverse	SACTAATCAATTTCCAAATCC	57.6	33.3
>11_R_Arthropoda_COI	Reverse	GACAAGCCAATTACCAAATC	59	40

>12_R_Arthropoda_COI	Reverse	KACAAGTCAATTTCCAAATCC	59.4	35.7
>13_R_Arthropoda_COI	Reverse	GGATAGTCAATTTCCAAATCC	59.3	38.1
>14_R_Arthropoda_COI	Reverse	SACTAATCAATTWCCAAATCC	56.4	33.3
>15_R_Arthropoda_COI	Reverse	KAMTAATCAATTACCAAATCC	53.8	28.6
>16_R_Arthropoda_COI	Reverse	TACAAGTCAATTWCCAAATCC	57.3	33.3
>17_R_Arthropoda_COI	Reverse	TACCAACCAATTTCCAAATCC	62.6	38.1
>18_R_Arthropoda_COI	Reverse	TACTAATCAATTTCCRAAKCC	57.6	33.3
>19_R_Arthropoda_COI	Reverse	WACTAATCAATTTCCAAAMCC	57	31
>20_R_Arthropoda_COI	Reverse	WACTAATCAATTWCCAAATCC	54.8	28.6
>21_R_Arthropoda_COI	Reverse	RAYTAATCAATTTCCAAATCC	56.8	28.6
>22_R_Arthropoda_COI	Reverse	RACTAATCAATTACCAAATCC	54.6	31

The primers were tested against DNA extractions from various arthropod orders to ensure that they would be amplify successfully. The primers were also tested against bat DNA, which was extracted from wing membrane. The wing membrane had been washed using ethanol, to avoid contamination with mite DNA which was found when using DNA extracted from fur. They were also tested against human DNA (taken from $CaCo^2$ cells at ~$2x10^5$) to ensure that the primers did not amplify these. All extractions were done using the QIAGEN DNeasy Blood and Tissue Kit and quantified using the Qubit HS Assay Kit and a Qubit Fluorometer. These underwent PCR using the reaction conditions detailed in section 4.3.1.2.

All Arthropoda PCR test reactions gave positive results, and all negative control reactions (bat and human) gave negative results.

In order to confirm that the new primer set will amplify a broader range of Arthropoda, a unique BLASTn database was created as follows. GI numbers were collected for all arthropod NCBI nucleotide entries using the taxid 6656. These GI numbers were used to create an Arthropoda only nt database using the command:

```
$ blastdb_aliastool -gilist sequence.gi.txt -db nt -out
nt_arthropoda -title nt_arthropoda -dbtype nucl
```

Each of the primers (from (Zeale et al., 2011), and from this study) were blast-searched against this database for exact matches (100% identity and a word length the same as the primer length). The primer sequences used in the search were non-degenerate and non-redundant.

```
$ blastn -query primers.fasta  -db nt_arthropoda -out
primers_arth_test.txt -task blastn-short -perc_identity 100 -
num_descriptions 50000 -num_alignments 1 -word_size 30
```

The number of exact matches for each was then counted. The Zeale primers had exact matches to 86,888 Arthropoda sequences, whereas the new primer set had 194,337 exact matches.

4.3.1.2. PCR replicates to avoid PCR bias

In order to avoid PCR bias, six PCR replicates were undertaken on each of the guano sample extraction.

10 *µ*l PCRs were prepared using a mixture of all of the primers shown in table 4.1, each at 5 *µ*M. Each PCR contained 1 *µ*l 5X Phusion High Fidelity buffer, 1 *µ*l of dNTPs at 2mM, 0.4 µl 50mM Mg++, 0.65 *µ*l forward primer mix, 0.65 *µ*l reverse primer mix, 0.05 *µ*l Phusion High Fidelity DNA polymerase, between 0.1-1 *µ*l of sample and 5.25-6.15 *µ*l ultrapure H_2O. Touchdown PCR was used in order to account for the differences in optimum annealing temperatures of the primers used (Don et al., 1991, Korbie and Mattick, 2008). Touchdown thermal cycling conditions were as follows: 5 mins at 94°C, followed by 20 cycles of 94°C for 30s, 58°C for 30s (decreasing by 0.3°C per cycle) and 72°C for 30s, followed by 5 cycles of 94°C for 30s, 52°C for 30s, then 72°C for 30s, followed by a final extension period of 72°C for 5 mins.

Reaction success was determined by 2% agarose gel electrophoresis stained with GelRed™ (Biotium, Inc.), and, if successful, were pooled, then purified and concentrated using homebrew SPRI beads (see appendix A.2. and (Rohland and Reich, 2012)). After purification, they

were eluted in 20 *μ*l of EBT buffer. EBT was used for elution rather than the more usual EB, as it increases the ease of bead migration during SPRI purification. The recipe for EBT can be found in the appendix A.3. After Purification, PCR replicates for each sample were pooled.

4.3.1.3. Other alterations to the library preparation protocol

The library preparation method was similar to that described in chapter three, however, due to the use of PCR products, fragmentation was not necessary, and the majority of the library preparation was undertaken in a post-PCR laboratory. The extractions and PCR set-up was done in a dedicated, PCR free, laboratory for working with modern DNA. PCR contamination is likely to be problematic as it could give confounding or false positive result. To further address this issue, all stages of extractions, PCR, and library preparation is accompanied by blank negative controls, see appendix A.1. These blank controls were pooled at the library preparation stage, assigned a uniquely indexed adapter, and sequenced along with the other samples.

Rather than Amplitaq Gold (Moretti et al., 1998), which was used for indexing PCR in the shotgun metagenome library preparation reactions, Phusion Hot Start High Fidelity was used (Lahr and Katz, 2009). This is because the carry-over of uracils was not required in the metabarcoding.

The rest of the library preparation protocol was as described in 3.3.1.4 (chapter 3). Largely, the same samples as used in chapter three (appendix E.3.4) were used in this chapter (appendix E.4.1). Each library was made using DNA from only one guano. Each species had 6 or 7 libraries (except for *M. alcathoe* which only has one). The indexes assigned were as follows;

Table 4.2: Indexes used in the indexing PCR and for de-multiplexing libraries in the metabarcoding Illumina MiSeq run

Species		D701	D702	D703	D704	D705	D706	D707	D708	D709
		ATTACTCG	TCCGGAGA	CGCTCATT	GAGATTCC	ATTCAGAA	GAATTCGT	CTGAAGCT	TAATGCGC	CGGCTATG
D501	TATAGCCT	*B. barbastellus*	*M. bechsteinii*	*M. daubentonii*	*M. nattereri*	*N. noctula*	*P. austriacus*	*P. pipistrellus*	*R. ferrumequinum*	*M. alcathoe*
D502	ATAGAGGC	*B. barbastellus*	*M. bechsteinii*	*M. daubentonii*	*M. nattereri*	*N. noctula*	*P. austriacus*	*P. pipistrellus*	*R. ferrumequinum*	*M. bechsteinii*
D503	CCTATCCT	*B. barbastellus*	*M. bechsteinii*	*M. daubentonii*	*M. nattereri*	*N. noctula*	*P. austriacus*	*P. pipistrellus*	*R. ferrumequinum*	*M. mystacinus*
D504	GGCTCTGA	*B. barbastellus*	*M. bechsteinii*	*M. daubentonii*	*M. nattereri*	*N. noctula*	*P. austriacus*	*P. pipistrellus*	*R. ferrumequinum*	*P. auritus*
D505	AGGCGAAG	*B. barbastellus*	*M. bechsteinii*	*M. daubentonii*	*M. nattereri*	*N. noctula*	*P. austriacus*	*P. pipistrellus*	*R. ferrumequinum*	*P. austriacus*
D506	TAATCTTA	*B. barbastellus*	*M. brandtii*	*M. mystacinus*	*N. leisleri*	*P. auritus*	*P. nathusii*	*P. pygmaeus*	*R. hipposideros*	*P. pipistrellus*
D508	CAGGACGT	*E. serotinus*	*M. brandtii*	*M. mystacinus*	*N. leisleri*	*P. auritus*	*P. nathusii*	*P. pygmaeus*	*R. hipposideros*	*P. pygmaeus*
D507	TGACTGAC	*E. serotinus*	*M. brandtii*	*M. mystacinus*	*N. leisleri*	*P. auritus*	*P. nathusii*	*P. pygmaeus*	*R. hipposideros*	*P. pygmaeus*
D509	GTCACATG	*E. serotinus*	*M. brandtii*	*M. mystacinus*	*N. leisleri*	*P. auritus*	*P. nathusii*	*P. pygmaeus*	*R. hipposideros*	*M. daubentonii*
D510	ACTGTACG	*E. serotinus*	*M. brandtii*	*M. mystacinus*	*N. leisleri*	*P. auritus*	*P. nathusii*	*P. pygmaeus*	*R. hipposideros*	*M. nattereri*
D511	GCGCATTC	*E. serotinus*	*M. brandtii*	*M. mystacinus*	*N. leisleri*	*P. auritus*	*P. nathusii*	Positive Control	*R. hipposideros*	*R. hipposideros*
D512	CTCTGGAA	*E. serotinus*	*M. daubentonii*	*M. nattereri*	*N. noctula*	*P. austriacus*	*P. pipistrellus*	*R. ferrumequinum*	Blank	*M. brandtii*

4.3.2. Bioinformatics

As the two MiSeq reads (forward and reverse) overlapped, just the forward reads were used. Data processing and analysis was as described in 3.3.2. The sample spreadsheet can be found in appendix B.2.

Primer sequences were removed from the FASTA sequences using fastx_trimmer from the FASTX-Toolkit (Gordon and Hannon, 2010) as follows;

```
$ fastx_trimmer -f 23 -l 189 -i [input_file.fasta] -o
[output_file.fasta]
```

All analysis was undertaken on data with and without primer trimming, in order to assess the impact of primer trimming. As the primer set was so degenerate, the impact of primer trimming was expected to be low.

4.3.2.1. Qiime analysis

The latest iBOl data for animals (COI) was downloaded (iBOL_phase_6.50_COI.tsv.zip) (Che et al., 2010). This was unzipped, and formatted for Qiime by the following command. The gb2qiime.py script was developed by Mike McCann, 2014 and was downloaded from https://groups.google.com/forum/#!topic/qiime-forum/O-U_DWRKOq0.

```
$ ./gb2qiime.py -i iBOL_phase_6.50_COI.tsv –s
iBOL_phase_6.50_COI_sequence.fasta -t
iBOL_phase_6.50_COI_taxonomy.txt -v 1
```

Macqiime version 1.9.1 (Caporaso et al., 2010) was installed, then opened using:

```
$ macqiime
```

Open reference OTU picking was performed on each sample using the following command.

```
$ pick_open_reference_otus.py -i [file.fasta] -o pick_otus -p
[parameter_file.txt]
```

The parameter file (parameter_file.txt) contained the following commands to specify forward and reverse strand matching, and Bray-Curtis beta diversity calculations (Bray and Curtis, 1957).

```
"pick_otus:enable_rev_strand_match True
pick_otus:max_accepts 1
pick_otus:max_rejects 8
pick_otus:stepwords 8
pick_otus:word_length 8
assign_taxonomy:reference_seqs_fp
[path_to]iBOL_phase_6.50_COI_sequence.fasta
assign_taxonomy:id_to_taxonomy
[path_to]iBOL_phase_6.50_COI_taxonomy.txt
beta_diversity:metrics bray_curtis"
```

The output files (biom format) were summarised into tab separated value tables, including the taxonomy information using the following:

```
$ biom convert -i [path_to]otu_table_mc2_w_tax.biom -o
[output.txt] --to-tsv --header-key taxonomy
```

4.3.3. Assessing the performance of the metabarcoding primers

The FASTA outputs were mined to determine which primers had been most successful at amplifying Arthropoda DNA and which primers caused the most spurious amplification (defined as any sequence which was not assigned as Arthropoda). The full (non-degenerate) primer sequences, and reverse complements of the sequences were identified and counted in the full FASTA output, in the subset assigned as Arthropoda, and the subset not defined as Arthropoda (table 4.4). Code can be found in appendix D.

4.3.4. Statistical analysis

Statistical analysis was undertaken as described in 3.3.3.

4.4. Results

The Illumina MiSeq run returned 12.89 GB of FASTQ data before adapter trimming and quality control, of which 1.78 MB was undetermined.

4.4.1. Hits before and after adapter removal and quality filtering

After quality trimming, adapter and primer removal, and conversion to FASTA format, there was a total of 2,618.602 MB of forward read data, with an average of 24.24 MB per library. Each species had an average 153 MB of data, with the negative control returning 0.602 MB, and the positive control returning 17 MB data (table 4.3). FastQC (Andrews, 2010) outputs for the dataset can be found in appendix C.4.1.

Figure 4.1. shows the breakdown of the sequence length distribution before cutadapt processing, and figure 4.2 shows the sequence length distribution after cutadapt processing (Martin, 2011).

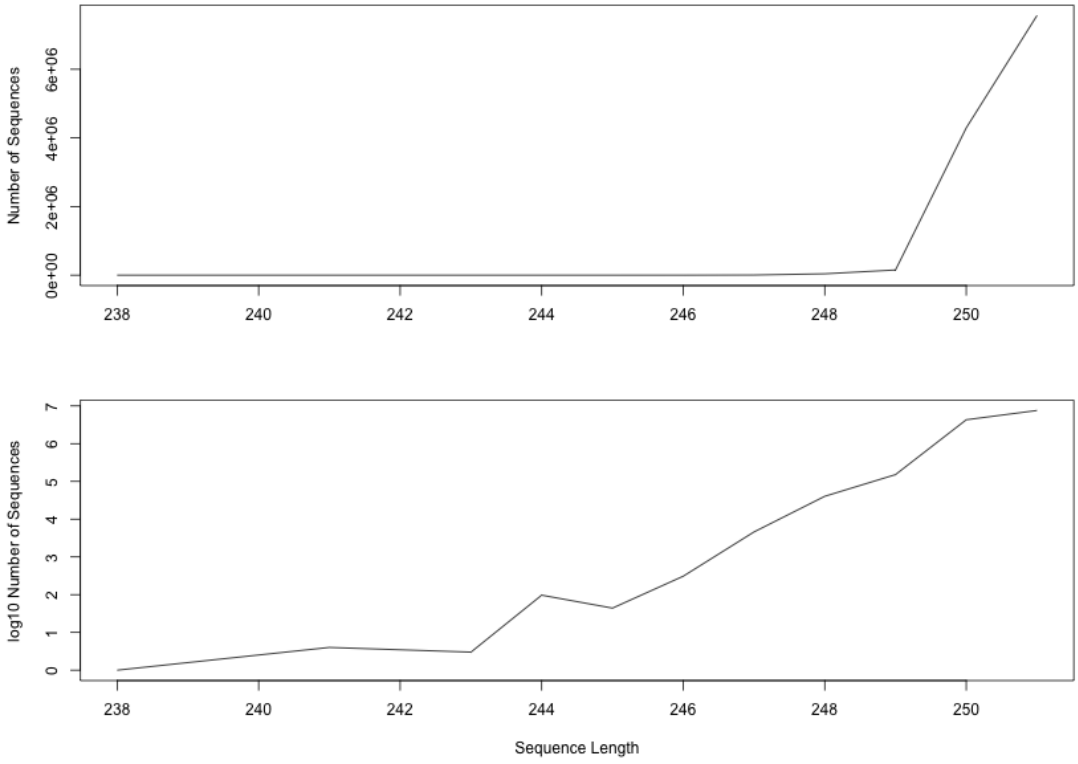


Figure 4.1: Sequence length distribution raw data

Before quality and adapter trimming after cutadapt. A. sequence length against the number of sequences, B. sequence length against the log10 number of sequences. Code in appendix D.2.6.

In figure 4.1 all of the reads have been sequenced to the end of the MiSeq's 250 cycles, due to incomplete adapter trimming in MiSeq Reporter. These are cleanly removed after adapter trimming, which reduces the sequences to 210 bp (this is the size of the 167bp amplicon plus primer sequences). After primer trimming, this sequence length distribution peak was reduced to 167bp (figure 4.2).

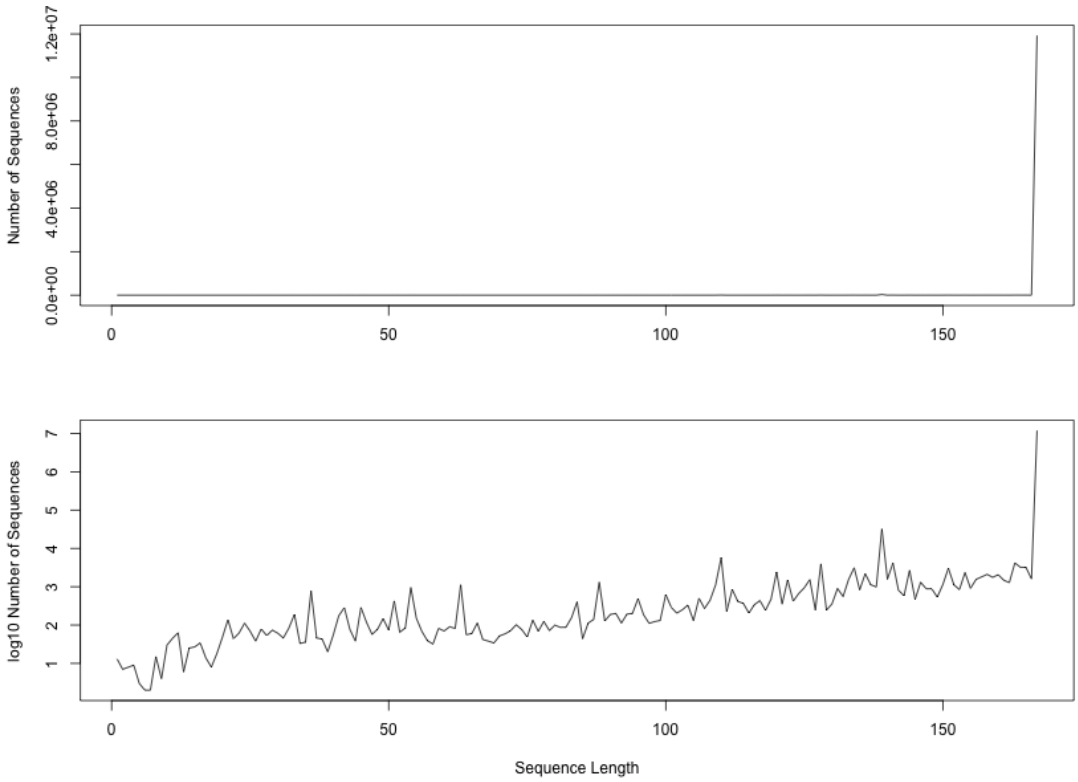


Figure 4.2: Sequence length distribution after quality, adapter, and primer trimming

Quality and adapter trimming undertaken using cutadapt, primer trimming used fastx_trimmer from the fastx toolkit. A. sequence length against the number of sequences, B. sequence length against the log10 number of sequences. Code in appendix D.2.6.

4.4.2. De-multiplexing sequence data

Table 4.3 shows the data assigned to each of the bat species, and the average data per library. *P. auritus* returned the largest dataset, at 277 MB of data, with an average of 39.57 MB per library, and *M. alcathoe* had the lowest data returned per library at 13 MB for the library created. Table 4.3 also shows the data assigned to each bat species and to the controls.

Table 4.3: Breakdown of the data returned for each species, and the data assignment

These data are after adapter and primer removal, quality trimming and blast score filtering.

Species	Count of libraries	MB data FASTA	Average MB/sample	Arthropoda Reads	Average Arthropoda reads/ sample
Blank	1	0.602	0.602	1,398	1,398
Positive control	1	17	17	69,166	69,166
B. barbastellus	6	104	17	431,817	71,970
E. serotinus	6	165	28	383,153	63,859
M. alcathoe	1	13	13	36,317	36,317
M. bechsteinii	6	103	17	237,396	39,566
M. brandtii	6	125	18	427,652	71,275
M. daubentonii	7	132	19	364,116	52,017
M. mystacinus	7	161	23	417,432	59,633
M. nattereri	7	150	21	542,432	77,490
N. leisleri	6	103	17	380,541	63,424
N. noctula	6	124	21	308,893	51,482
P. auritus	7	277	40	1,151,665	164,524
P. austriacus	7	246	35	906,259	129,466
P. nathusii	6	163	27	672,403	112,067
P. pipistrellus	7	191	27	776,673	110,953
P. pygmaeus	7	223	32	626,102	89,443
R. ferrumequinum	6	189	32	424,970	70,828
R. hipposideros	7	132	19	445,564	63,652

Generally, the bat species with larger datasets had a larger number of sequences assigned to Arthropoda (figure 4.3). This was a fairly strong trend- the R^2 for the amount of data returned against the number of Arthropoda BLAST hits was 0.841. The point with the lowest data returned (far left) was the blank. Despite all of the species having similar numbers of samples processed, there was variability in the number of reads returned for each species.

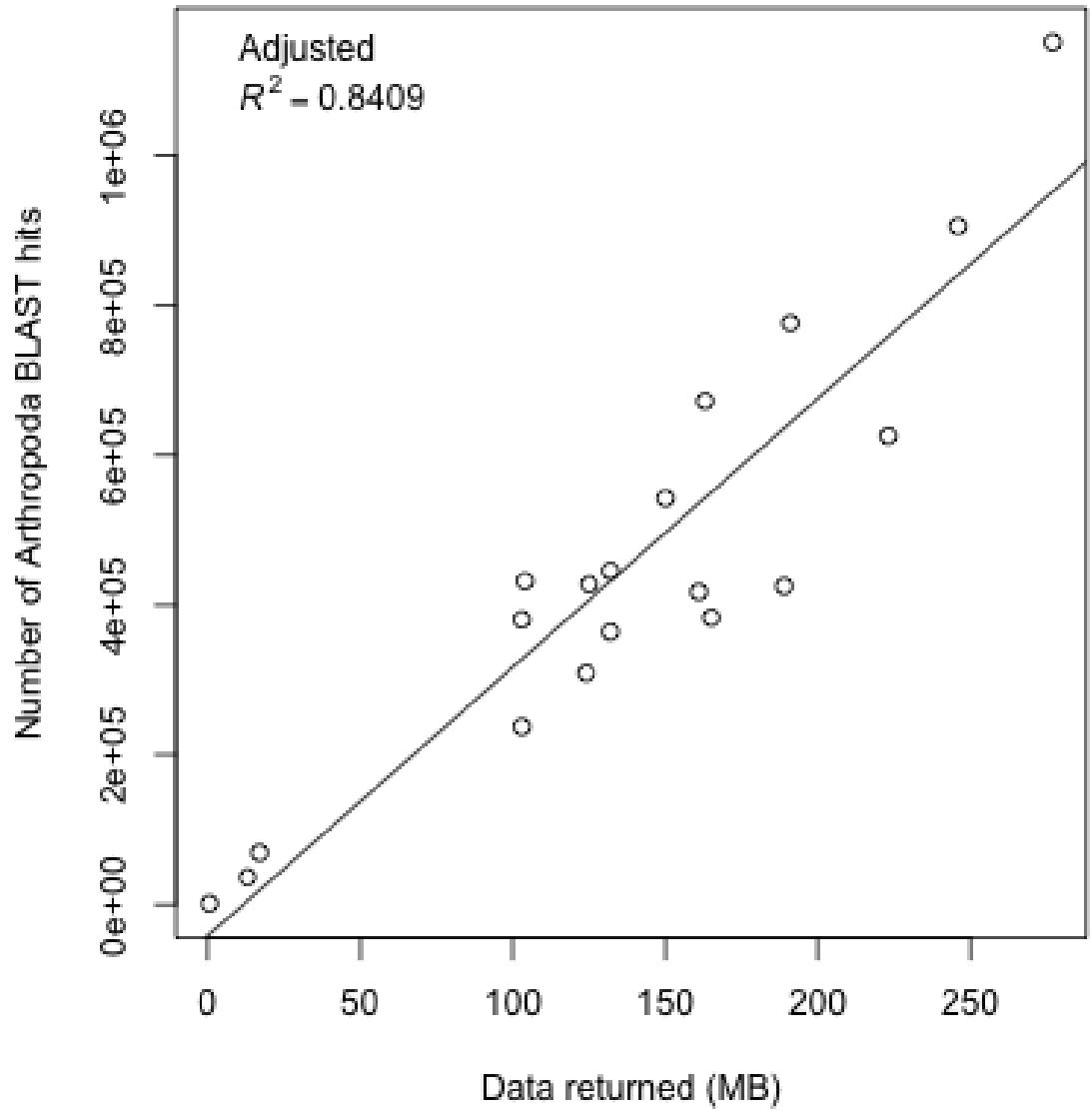


Figure 4.3: The data (MB) returned for each of the species datasets against the number of Arthropoda BLAST hits

Code can be found in appendix D.2.7.

The reason that the R^2 value in figure 4.3 was not closer to 1 is that ~9.845% of the data returned was not assigned to Arthropoda using MEGAN5 (figure 4.4). This was due to spurious amplification of other components of the guano, by the primers. Additionally, 69,615 reads (0.638%) of the total dataset were unassigned.

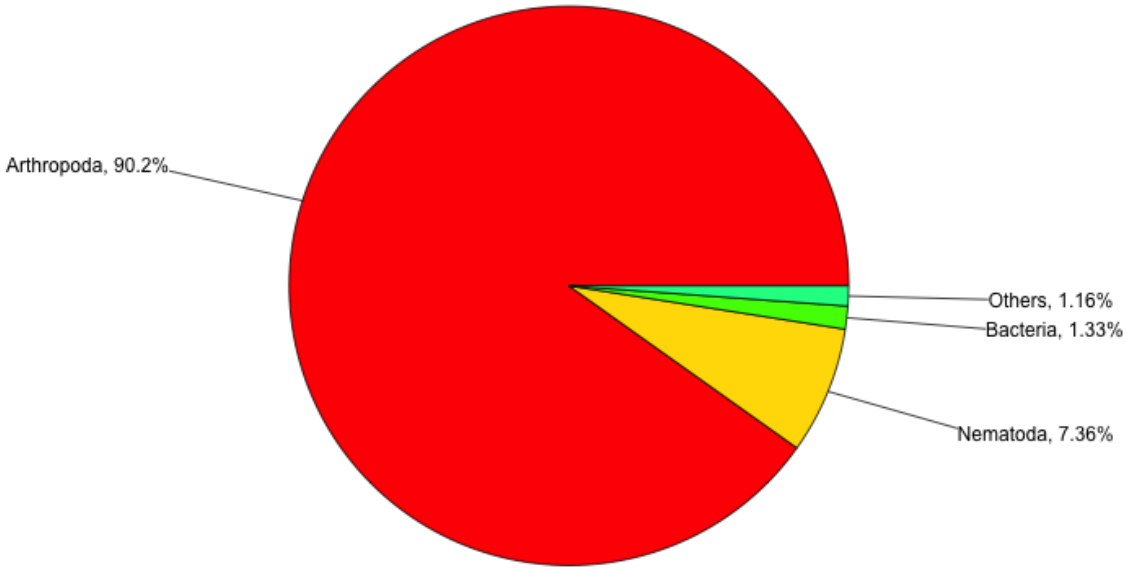


Figure 4.4: Overview of the assignment of the Illumina metabarcoding sequencing data for all samples

"Other" is comprised of Fungi (0.53%), Chordata (0.44%), Lophotrochozoa (0.09%), Stramenopiles (0.06%), Streptophyta (0.02%), Metazoa (0.008%), Amebozoa (0.0007%), Cnidaria (0.0006%), Rhodophyta (0.0004%), Viruses (0.0002%), and Scalidophora (0.0001%). Code in appendix D.2.8.

90.2% of the full dataset was assigned to Arthropoda (figure 4.4). The spurious data (all data not assigned to Arthropoda) was discarded from dietary analyses. Table 4.4 shows the how many of the primer sequences (section 4.3.3) were found in the Arthropoda datasets (before primer trimming), which were identified in the spurious amplifications (defined here as any sequence not assigned to Arthropoda).

Table 4.4: The representation of each primer sequence

Within the full dataset, in the subset containing only Arthropoda reads, and in the subset containing only non-Arthropoda reads. Counts are sum of 5'-3' sequences and the reverse complement.

Primer name	Total hit count	% assigned as non-Arthropoda	% assigned as Arthropoda
>1_F_Arthropoda_COI	942,211	1.08	98.92
>2_F_Arthropoda_COI	2,798,504	0.46	99.54
>3_F_Arthropoda_COI	153,326	0.32	99.68
>4_F_Arthropoda_COI	143,528	0.48	99.52

>5_F_Arthropoda_COI	4,064,892	0.27	99.73
>6_F_Arthropoda_COI	13,104	0.89	99.11
>7_F_Arthropoda_COI	2,844,699	0.76	99.24
>8_F_Arthropoda_COI	772,638	0.26	99.74
>9_F_Arthropoda_COI	671,501	0.28	99.72
>10_F_Arthropoda_COI	2,829,76	0.32	99.68
>11_F_Arthropoda_COI	66,010	0.49	99.51
>12_F_Arthropoda_COI	1,328,380	0.56	99.44
>13_F_Arthropoda_COI	2,060,389	0.28	99.72
>14_F_Arthropoda_COI	1,320,923	0.26	99.74
>15_F_Arthropoda_COI	17,263	42.83	57.17
>16_F_Arthropoda_COI	932,421	0.25	99.75
>17_F_Arthropoda_COI	1,260,774	0.37	99.63
>18_F_Arthropoda_COI	72	42.77	57.23
>19_F_Arthropoda_COI	4,468,413	0.27	99.73
>1_R_Arthropoda_COI	855,272	26.65	73.35
>2_R_Arthropoda_COI	1,991,962	9.71	90.29
>3_R_Arthropoda_COI	3,022,747	8.06	91.94
>4_R_Arthropoda_COI	3,057,691	7.02	92.98
>5_R_Arthropoda_COI	517,068	8.79	91.21
>6_R_Arthropoda_COI	234,694	19.00	81.00
>7_R_Arthropoda_COI	181,984	13.55	86.45
>8_R_Arthropoda_COI	92,844	10.48	89.52
>9_R_Arthropoda_COI	357,306	14.37	85.63
>10_R_Arthropoda_COI	2,181,935	6.73	93.27
>11_R_Arthropoda_COI	186,718	90.16	9.84
>12_R_Arthropoda_COI	1,143,332	15.05	84.95
>13_R_Arthropoda_COI	549,083	17.10	82.90
>14_R_Arthropoda_COI	2,650,304	7.25	92.75
>15_R_Arthropoda_COI	592,267	9.62	90.38
>16_R_Arthropoda_COI	894,605	17.64	82.36
>17_R_Arthropoda_COI	160,328	35.86	64.14
>18_R_Arthropoda_COI	1,209,923	21.40	78.60
>19_R_Arthropoda_COI	262,715	12.42	87.58
>20_R_Arthropoda_COI	3,465,596	7.43	92.57
>21_R_Arthropoda_COI	3,593,872	7.55	92.45
>22_R_Arthropoda_COI	647,148	8.44	91.56

Overall, 90.82% of primer hits were found in sequences assigned to Arthropoda, with others likely from spurious amplification, or as carryover from kit contamination (see section 3.4.4.1). The most successful primer was >19_F_Arthropoda_COI which was found in 4,468,413 sequences. The least successful primer was >18_F_Arthropoda_COI, which was only found in 72 reads out of the total dataset. The primer with the highest proportion of Arthropoda sequence hits was primer >16_F_Arthropoda_COI, which had 99.75% of its hits in sequences assigned to Arthropoda. Primer >11_R_Arthropoda_COI had the highest proportion of hits that were spurious sequences.

The proportion of kit derived non-target reads can be reduced by increasing the number of PCR cycles, as they get swamped by the PCR products, however, increasing the number of cycles would increase the chance of introducing PCR bias and spurious amplifications. There are two major types of PCR bias (Acinas et al., 2005): the first is PCR bias caused by the unequal amplification of target molecules (Polz and Cavanaugh, 1998). The second is sequence artefacts caused by PCR errors. This may be the formation of chimeric (Brakenhoff et al., 1991) or heteroduplex molecules (Ruano and Kidd, 1992), or caused by DNA polymerase error (Kobayashi et al., 1999). It is therefore unadvisable to use more PCR cycles than necessary, and the resulting non-target (e.g. non-Arthropoda) DNA can simply be discarded.

4.4.3. Control analysis

4.4.3.1. Negative control

The blank library forward read returned 0.602 MB of FASTA data, with a total of 1,757 assigned sequences. There were Arthropoda reads, however the blank library had a very low number of reads in comparison to the number of reads returned for each of the libraries. The average number of Arthropoda FASTA reads from all the sample libraries was 77,618 (table 4.3). It is possible that PCR amplification of DNA was more successful than some of the other libraries due to the lack of PCR inhibitors, which commonly occur in guano (Idaghdour et al., 2003,

Puechmaille et al., 2007, Taberlet et al., 1999). Additionally, where all of the other libraries were diluted to 4nM before sequencing, the blank library was not diluted (as it was at <4nM). Figure 3.5 shows the breakdown of the sequences returned. There were no bat sequences returned, indicating that there has been no contamination of the reagents during the DNA extraction. Many of the sequences returned from the blank library likely originated from the reagents (Salter et al., 2014a, Salter et al., 2014b).

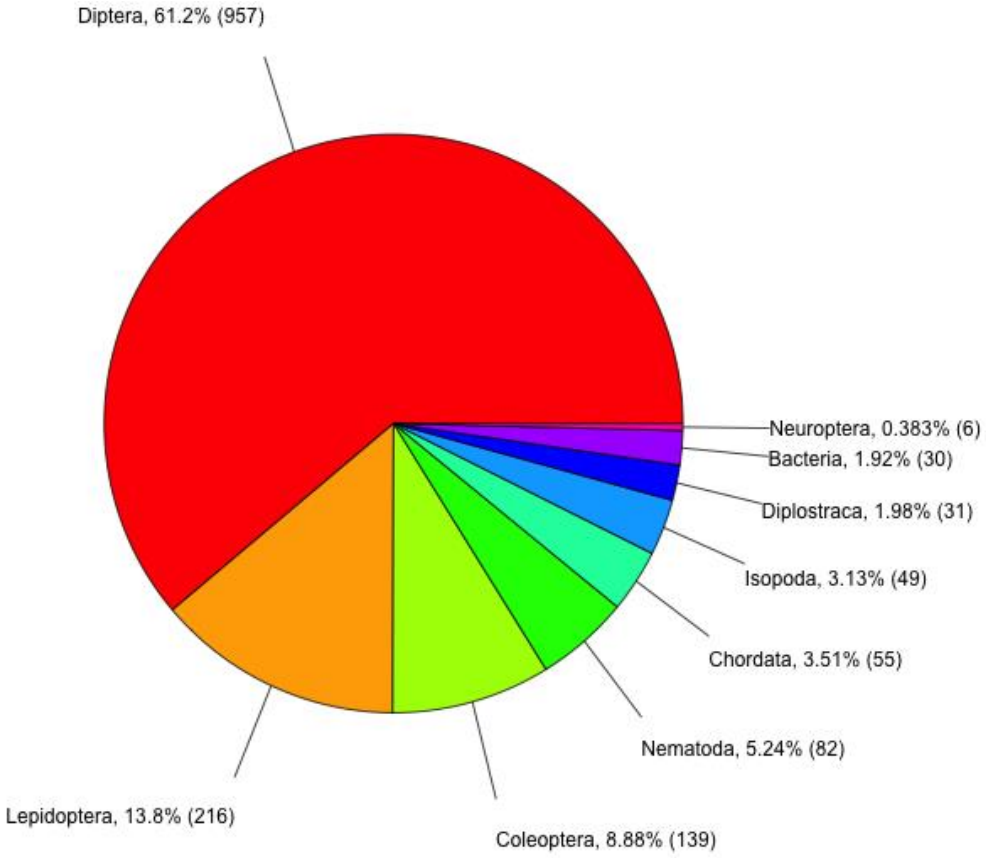


Figure 4.5: The taxonomic profile of the blank library

Showing the number of sequence hits in brackets. Code in appendix D.2.8.

4.4.3.2. Positive control

The positive control sample was from a hand reared *P. pygmaeus*, which had been feed on *Tenebrio molitor* (meal worm) larval instars, which belong to the order Coleoptera. This is reflected in the data with 51,171 reads (68.3%) assigned to Coleoptera by MEGAN5 (figure 4.6). Of these 51,096 reads were assigned to *T. molitor* at species level.

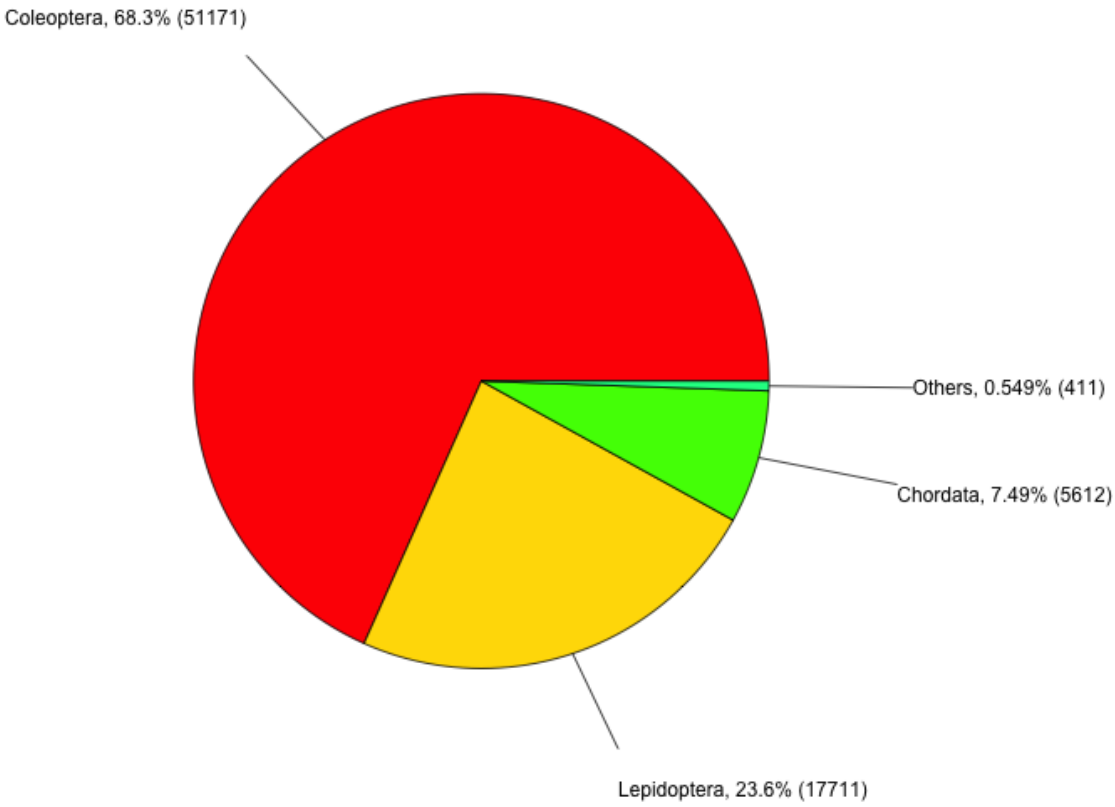


Figure 4.6: The taxonomic profile of the positive control library
The positive control library was from a captive bred *P. pygmaeus.* Showing the number of sequence hits in brackets. Code in appendix D.2.8.

22.6% of the positive control library was assigned to Lepidoptera. Of the 17,711 reads, 16,826 reads were assigned to *Galleria mellonella*, the greater wax moth. The larvae of *G. mellonella,* the waxworm, is a commercially bred for scientific research (Aperis et al., 2007, Kavanagh and Reeves, 2004), and for bait (Klingen et al., 2002) and pet food (Finke, 2002), suggesting that despite our records, the captive bat was fed on *G. mellonella* as well as *T. molitor*.

4.4.4. Rarefaction analysis

Figure 4.7. shows the rarefaction plot for the datasets of each bat species at order level. All of the datasets plateau quickly and have levelled off well.

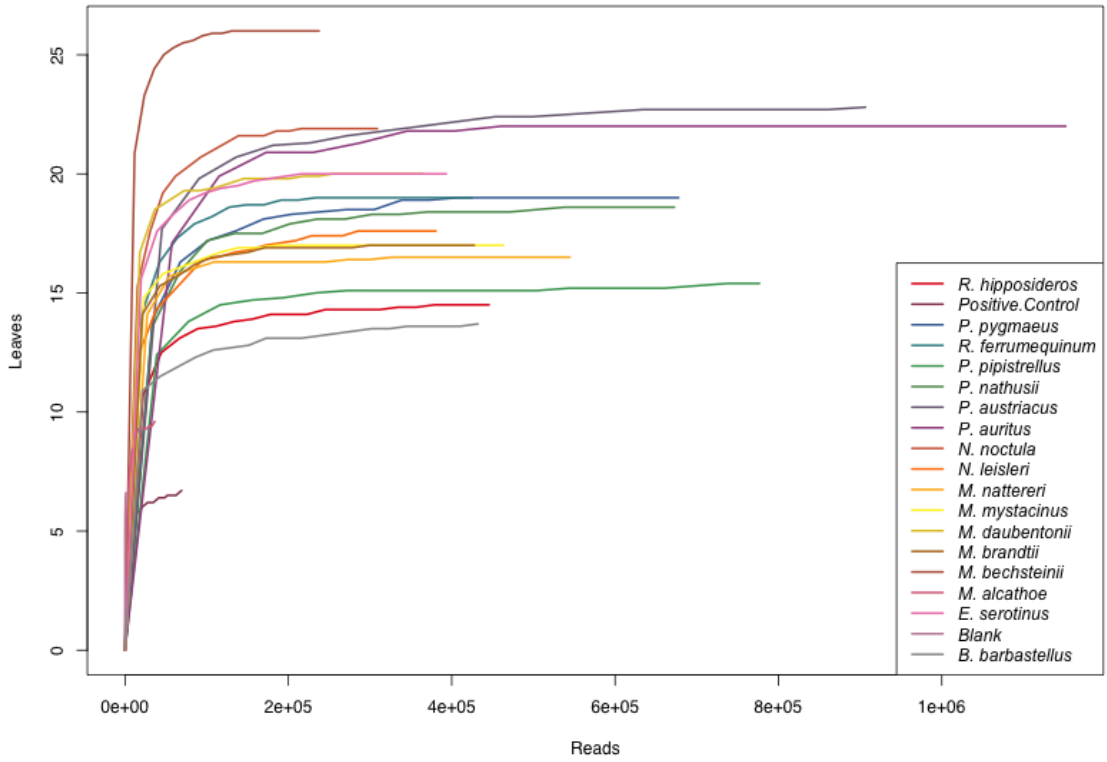


Figure 4.7: Rarefaction plot for the Arthropoda metabarcoding dataset at order level

Excluding controls. Code in D.2.9.

The alcathoe, which has a smaller dataset, is less clear. This is probably an artefact of the lack of technical replicates. However, it does level off, just not as clearly as the other datasets. This suggests that the sampling effort has been sufficient for identifying the majority of the taxon present across the full dataset (Chroňáková et al., 2009, Heck Jr et al., 1975).

4.4.5. Dietary composition of guano

The major source for arthropod DNA from guano would most likely from dietary origins. 90.2% of the full dataset was assigned to Arthropoda (figure 4.4). Reads were assigned to 38 orders of Arthropoda, and to 634 species. Assignment to a taxa required >5 sequences assigned. Figure 4.8 shows the breakdown of the arthropod orders seen across the full dataset.

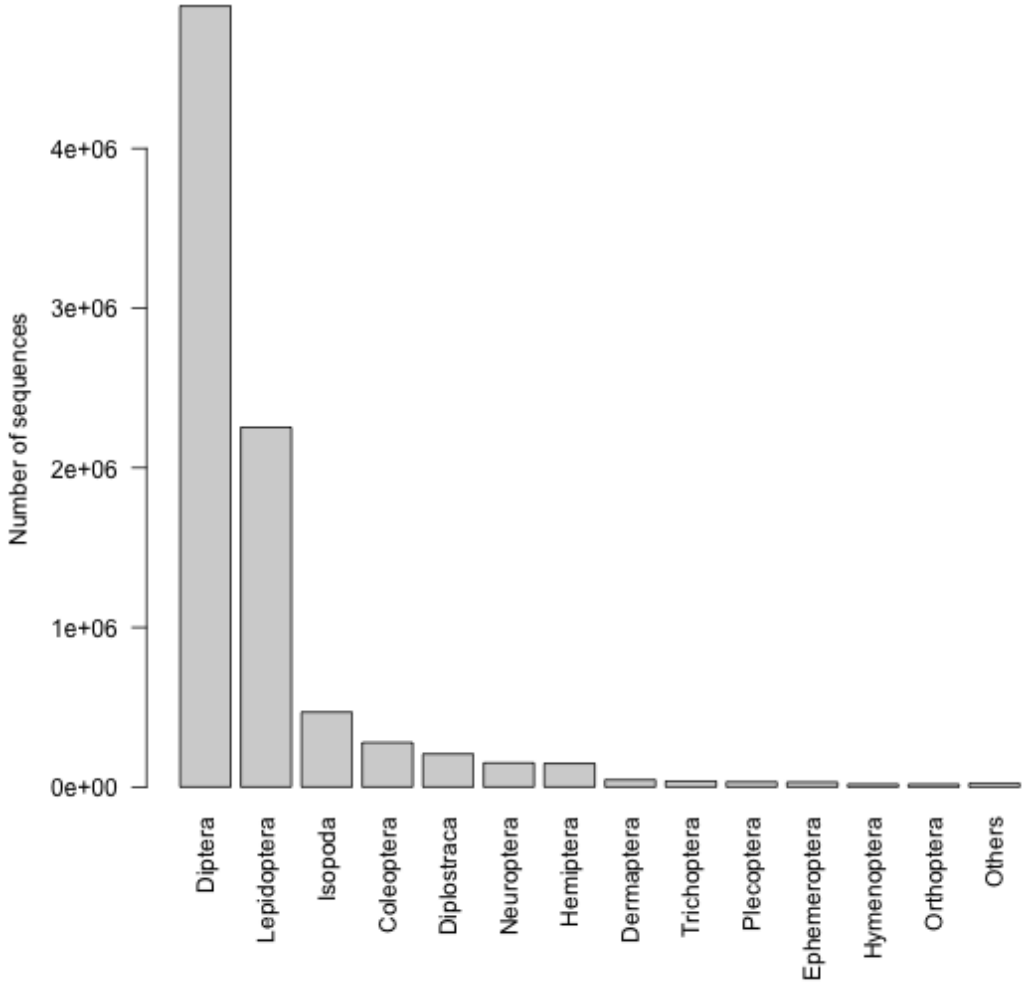


Figure 4.8: Assignment of Arthropoda sequences for all samples
These are shown to the level of order. Code in appendix D.2.10.

As was seen in the shotgun metagenome analysis (chapter 3) and previous studies (see chapters one and two), Lepidoptera and Diptera comprise the majority of bat diet species (figure 4.8). The largest order is the Diptera, with 4,892,274 reads in the full dataset. The second largest order, the Lepidoptera, has 2,252,856 reads.

Figure 4.9 shows the variation in diets of all of the bat species collapsed to order level. The major distinguisher between dietary types is a preference for either Diptera or Lepidoptera.

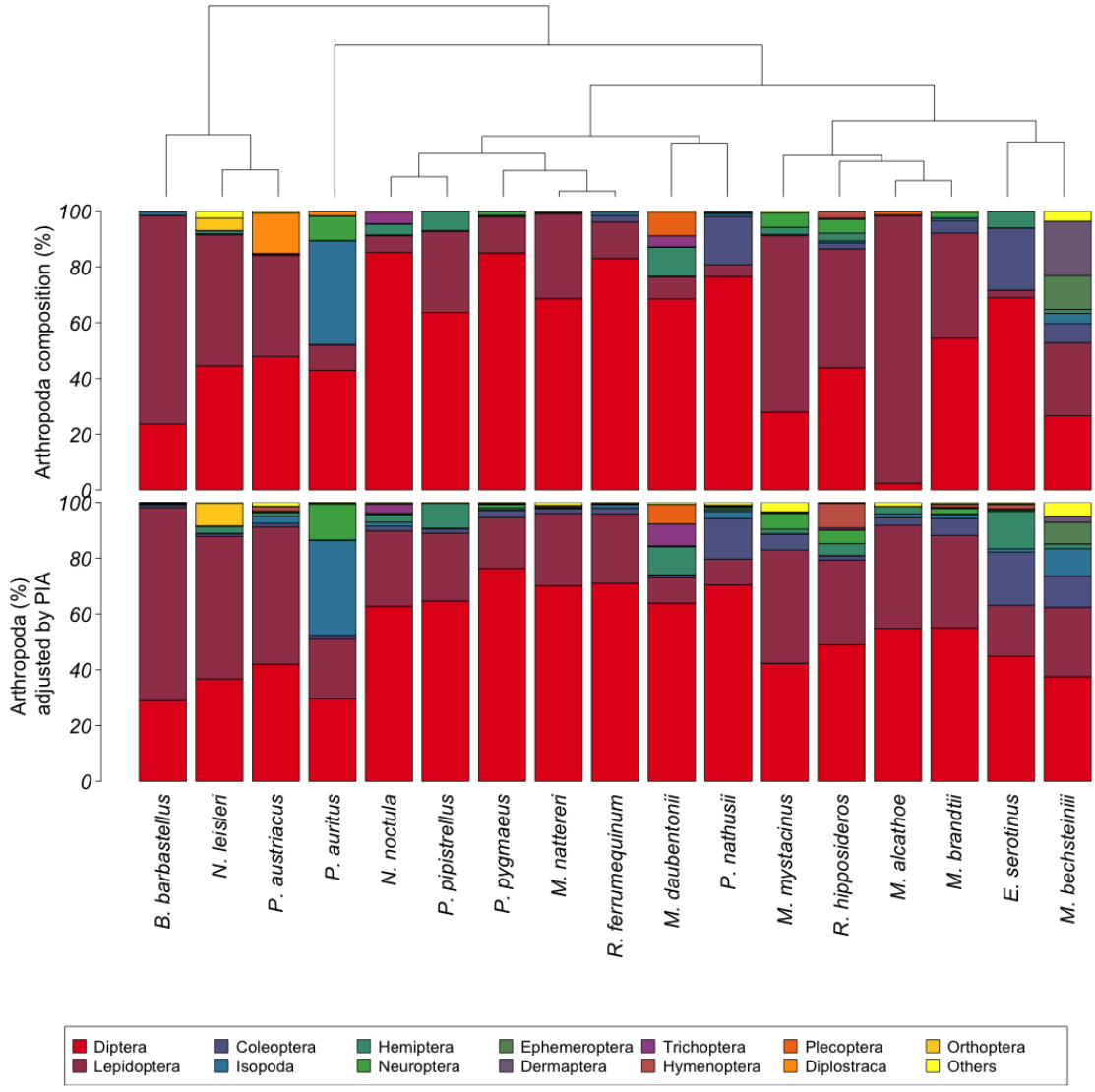


Figure 4.9: The diet of each bat species at order level a. before and b. after PIA
Prey taxa have been grouped by order. Diets sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2.

4.4.6. Validation of the diet data

In order to validate the analysis conducted using BLAST and MEGAN5, PIA (section 3.3.2.5.1), Qiime analysis and genome size adjustment (section 3.3.2.6) was undertaken on the metabarcoding data.

4.4.6.1. PIA analysis of the data

PIA analysis of a subset of 500 reads per sample of the Arthropoda data largely supports the breakdown of the bat diets (table 4.5), although it resulted in the discarding of 33.57% of the 54,000 Arthropoda reads

tested. These data will be compared to the literature dataset and the metagenome dataset in chapter five.

Table 4.5: Data assignment before and after phylogenetic intersection analysis

Species	Number of Arthropoda BLAST hits before PIA	Number of Arthropod BLAST hits after PIA
B. barbastellus	3000	2137
Blank	500	248
E. serotinus	3000	1653
M. alcathoe	500	114
M. bechsteinii	3000	1723
M. brandtii	3500	1911
M. daubentonii	3500	2720
M. mystacinus	3500	1873
M. nattereri	3500	2364
N. leisleri	3000	2240
N. noctula	3000	2310
P. auritus	3500	2896
P. austriacus	3500	2288
P. nathusii	3000	2285
P. pipistrellus	3500	2143
P. pygmaeus	4000	2464
Positive control	500	443
R. ferrumequinum	3000	1876
R. hipposideros	3500	2183

The positive control library had the highest proportion of reads retained after PIA (88.6%), and *M. alcathoe* the highest level discarded (22.8% retained).

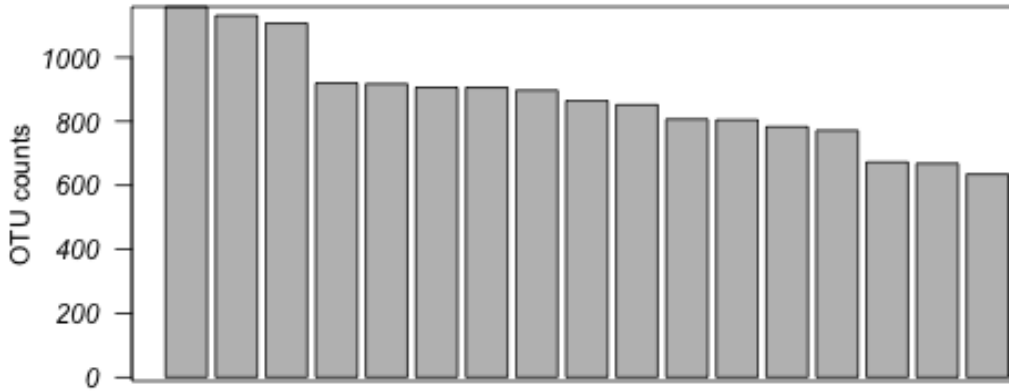
In the full dataset there were 38 Arthropoda orders identified, of which 24 were retained after Phylogenetic Intersection Analysis (figure 4.9). The orders discarded were typically those with low representation within the datasets: the majority of highly represented orders were not dramatically reduced.

It is important to note that the positive control data after PIA was highly similar to the data before PIA (figure 4.6). Despite 215 reads being discarded, Coleoptera was still the highest represented taxa (73.98% before, 56.75% after PIA) and Lepidoptera the second (25.60% before, 33.24% after PIA) with 3.33 assigned to other Taxa after PIA.

4.4.6.2. Qiime analysis

Qiime (Caporaso et al., 2010) analysis was undertaken using the data downloaded from the Barcode on Life Database (section 4.3.2.1) (Che et al., 2010, Ratnasingham and Hebert, 2007). 18,999 OTUs were identified in the full dataset, and 4,989 of these were identified at order level, and assigned to 24 orders. Sample BatID 899, a *R. ferrumequinum,* had the highest number of OTUs (2,765). The blank library had the lowest number of OTUs (78 OTUs). The sample with the lowest number of OTUs was the *E. serotinus* sample BatID 2297, with 425 OTUs. *P. auritus* had the highest average number of OTUs (1,158), and *M. alcathoe* was the species with the lowest number of OTUs (635) (figure 4.10).

a)



b)

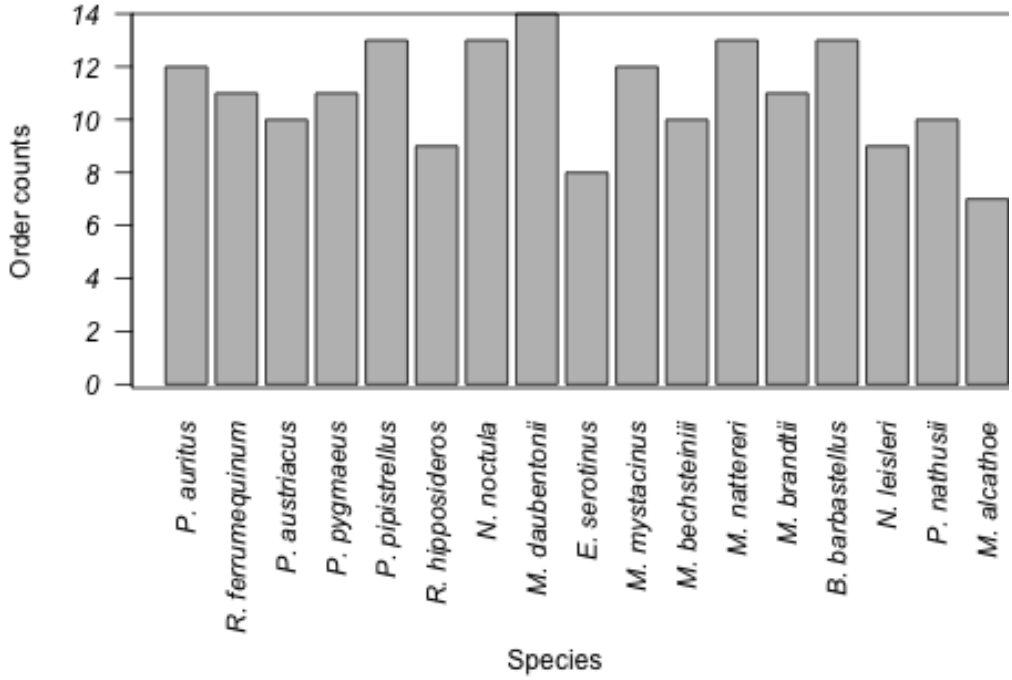


Figure 4.10: The average number of a) OTUs and b) orders identified in the diets of each bat species

As reported by Qiime. Code adapted from appendix D.2.10.

Generally, the assignment to the major dietary Arthropoda orders reported in Qiime are similar to those seen previously (figures 4.9 and 4.11), however, large proportion (73.74%) of the OTUs were unassigned to Arthropoda orders. As a result of the large number of OTUs discarded,

the species with the highest number of orders seen (*M. daubentonii,* 14 orders), did not have the highest number of OTUs assigned.

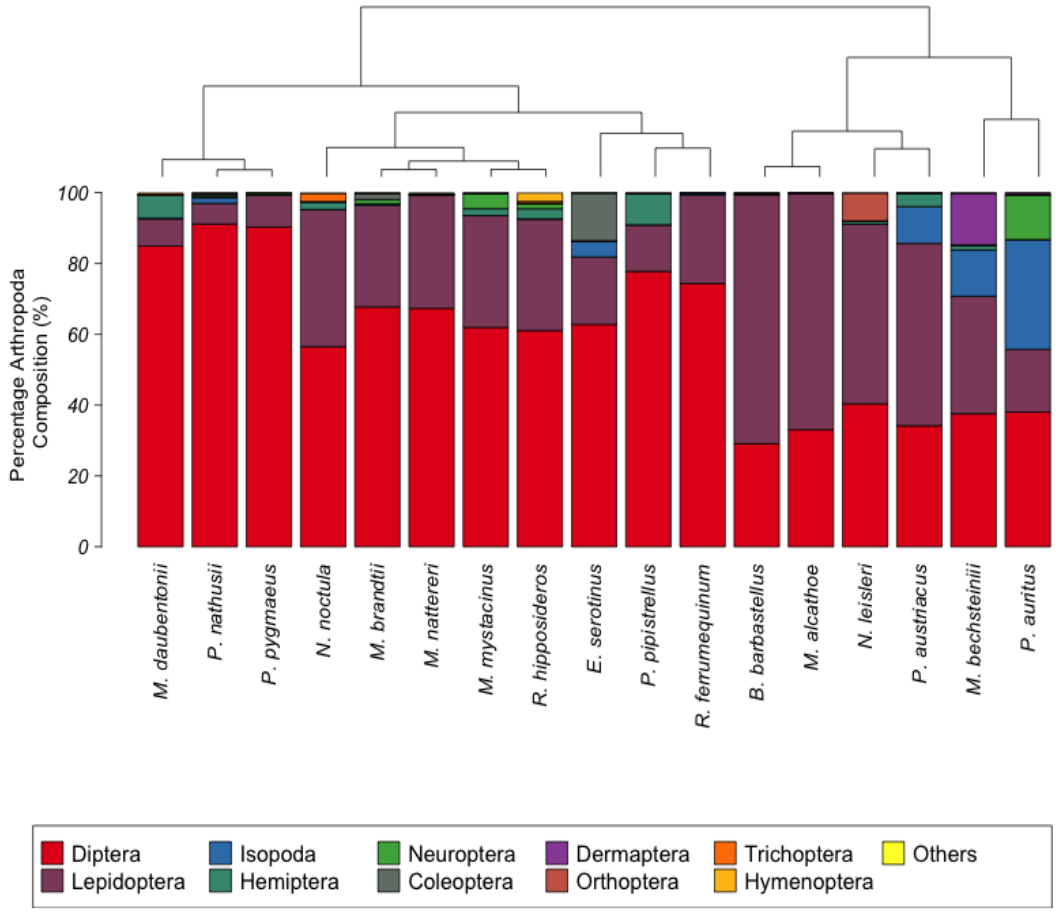


Figure 4.11: The diet of each bat species at order level, as reported by Qiime
The average diets of each of the prey species. Prey taxa have been grouped by order. Diets sorted using complete-linkage clustering using R hclust package. Code adapted from appendix D.2.2.

Some orders such as Coleoptera are poorly represented in the Qiime data (figure 4.11) when compared to the data from our previous analyses (figure 4.9). This was particularly striking in the data for the positive control. This had been feed on *T. molitor* and possibly on *G. mellonella* (section 4.4.3.2). In the BLAST and MEGAN5 analysis, 75.09% of the data was assigned to *T. molitor* (Coleoptera) and 24.73% to *G. mellonella* (Lepidoptera). However, in the Qiime and BOLD analysis, 99.28% of the Arthropoda data was assigned to Lepidoptera, 0.668% to Diptera, and only 0.295% to Coleoptera. In this sample 23.29% of the data was

assigned to Arthropoda, and 76.71% of the data was unassigned. This may be as a result of poor coverage of the Coleoptera in the Barcode on Life Database. This may be the cause of the lower than expected Coleopteran data in the molecular data reported in (Razgour et al., 2011). The blank library had a small number of reads assigned to Arthropoda OTUs (1,093 reads), the majority of which were Diptera, as was seen in section 4.4.3.1.

4.4.6.3. The importance of removing primer sequences from metabarcoding data

Primer sequences are trimmed from DNA sequences before analysis as they can mask sequence diversity in the priming regions caught by mismatch hybridisation. Whilst there was little difference in the overall data assignments before and after primer trimming, there was a slightly higher diversity seen at order level in each of the methods of analysis (BLAST and Qiime). Additionally, after PIA, a lower proportion of reads were discarded in the primers trimmed dataset (33.57%) than in the data with primers intact (41.2%) This was likely due to the high level of degeneracy in the primers, with the most similar primers binding to the target sequences, thus avoiding much of the mismatch hybridisation. This suggests that primer diversity is not sufficient to reflect taxon diversity. All figures and data presented in this chapter are after primer removal.

4.4.7. Dietary diversity, niche breadth, and overlap

Figure 4.12 shows the dietary diversity (Shannon-Weaver's diversity index, grey bars), and the niche breadth (Levin's index, black bars). These were calculated using all data assigned at order level (discounting anything assigned only to class or phylum, and collapsing to order anything assigned at family, genus, or species). As discussed in section 7.1.6, each species is based on 6 or 7 samples (except *M. alcathoe*), and as a result does not likely represent the true dietary diversity and niche breadths of each species (it is likely an underestimate). Future work should include an increased number of samples to allow robust statistical analysis.

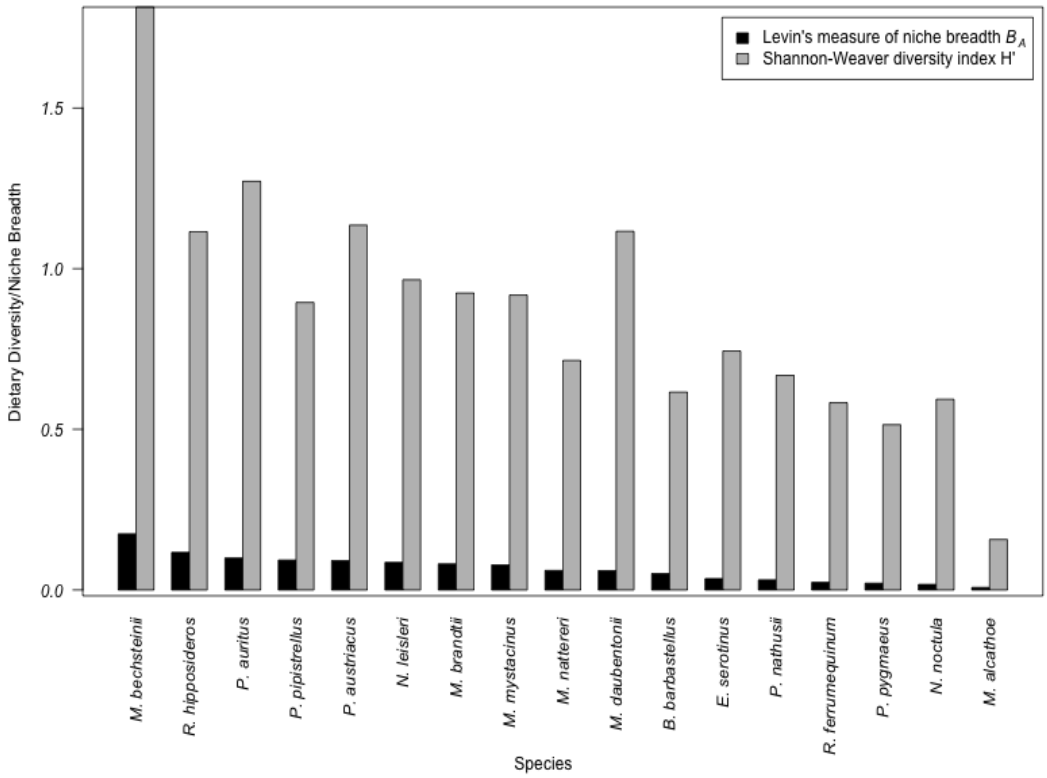


Figure 4.12: Dietary diversity and niche breadth of each bat species

The dietary diversity calculated using Shannon-Weaver diversity index (H) (grey) and niche breadth calculated using Levins standardised index (B_A) (black). Arthropod data assigned at order level. Code adapted from appendix D.2.3. For n-values, see E.4.1.

M. alcathoe with a Shannon-Weaver diversity index of 0.16, has the smallest diversity, whereas *M. bechsteinii* with an index of 1.81, has the greatest diversity. The species with the smallest niche is also *M. alcathoe,* with an adjusted Levin's measure of 0.007, whereas the species with the broadest niche is also *M. bechsteinii* with a measure of 0.17.

The Shannon-Weaver index is sensitive to rare prey species, whereas Levins's measure gives weighting to the dominant prey species. The Shannon-Weaver diversity is much higher than in the metagenome dataset due to the higher return of taxa, however, the diversity is still lower. This suggests that most bats feed primarily on few arthropod taxa, but occasionally feed on other taxa.

As was seen in the shotgun metagenome data (chapter 3) and in the literature review (chapter two), the Lepidoptera and Diptera form the basis of the diets of the Great British bats studied (figures 4.8, and 4.9). This allows the formation of three broad dietary guilds (figure 4.13). The first is defined by high proportions of Diptera (40-65%) and of Lepidoptera (30-50%). This guild contains *M. nattereri*, *R. hipposideros, M. mystacinus, P. pipistrellus, P. austriacus, M, brandtii*, and *N. leisleri.* The second guild contains all of which have a large proportion (>65%) of their diets comprised of Diptera, and do not feed on large proportions of Lepidoptera (<13%). The third guild includes *B. barbastellus* and *M. alcathoe* which feed almost solely on Lepidoptera (77% and 96% respectively).

P. auritus and *M. bechsteinii* do not cluster well into any guild, but are most similar to guilds 2 and 1 respectively. These bat species have the high Shannon-Weaver diversity indexes (H') and Levin's measure of niche breadth (B_A), and thus should be considered generalist species (section 4.5.2).

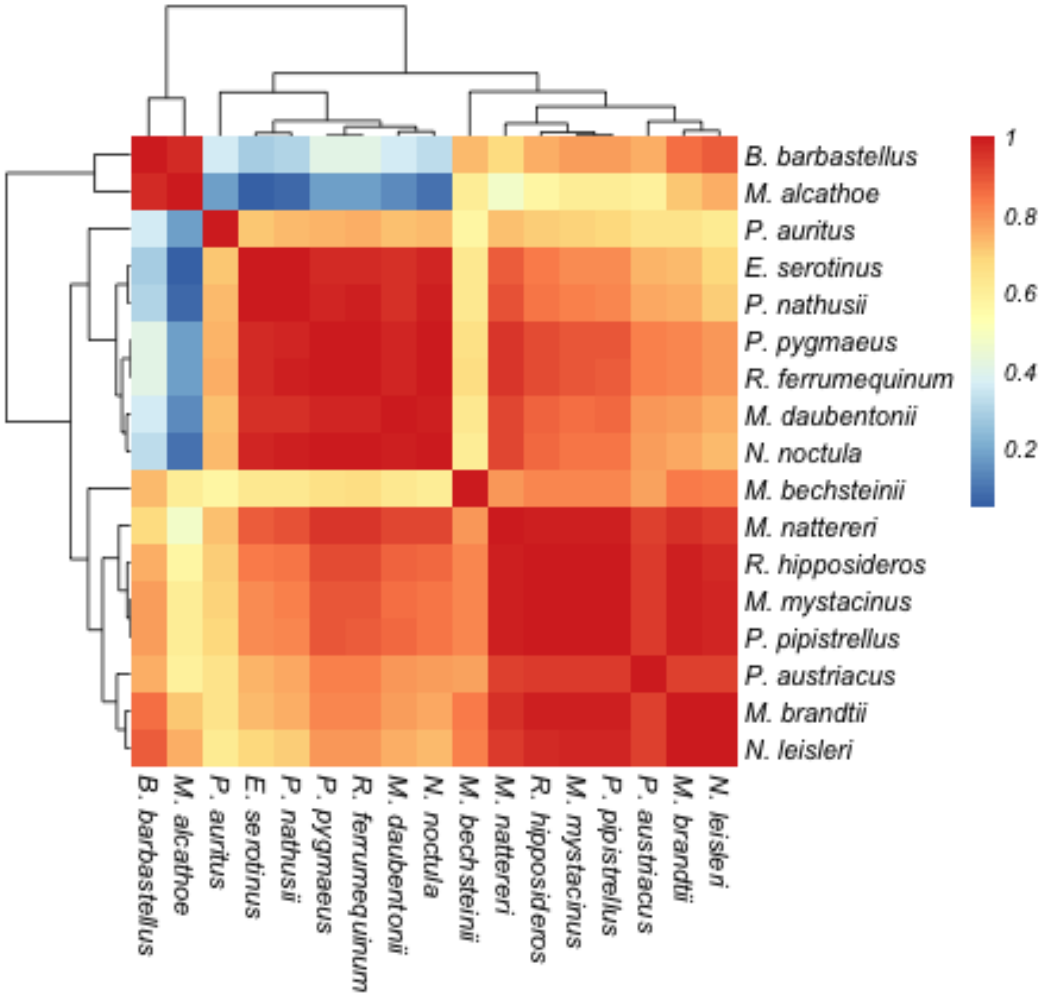


Figure 4.13: The dietary overlap of each bat species with Arthropoda at order level Calculated using Pianka's index of niche overlap. A value of 1 (red) suggests that the diets are identical, whilst a value of 0 (blue) indicates that there is no overlap. Species have been hierarchically clustered by niche overlap similarity. Code adapted from appendix D.2.4.

The mean overlap was 0.779, with the highest overlap (0.999) being between *R. ferrumequinum* and *P. pygmaeus*, and the lowest overlap (0.047) between *E. serotinus* and *M. alcathoe*. As was seen in chapter 3 (section 3.4.10), *P. pygmaeus* appears to be in direct competition with both *R. ferrumequinum* and *P. nathusii*. In this dataset, it also appears to be in competition with *E. serotinus, M. daubentonii* and *N. noctula.* Combined with it's low dietary diversity and niche breadth, *P. pygmaeus* appears to have a high vulnerability to extinction pressures (Boyles and Storm, 2007).

4.5. Discussion

4.5.1. Guano quality

It was possible to amplify the 167bp amplicons for all samples, indicating reasonable DNA quality. For some samples, the PCR replicates had to be repeated to ensure that there were 6 successful replicates for each sample. In some cases, the DNA extracts required dilution before PCR, in-order to achieve PCR success. This is likely a result of the presence of PCR inhibitors.

DNA from coprophageous arthropods, including the Calliphoridae, and Muscida (Galante and Marcos-Garcia, 2004, Waage, 1979) may be indicators of Arthropoda feeding on guano post defecation. DNA from the Calliphoridae was identified in all of the species datasets, with >250,000 reads assigned to the genus Pollenia originating in the BatID sample 1701. This *P. auritus* sample was collected in north Somerset in the autumn of 2012. Pollenia includes the cluster flies, which overwinter in attics and lofts (Greenberg, 1998).

4.5.2. Variation in diet between bat species

The guild structure has no apparent reflection of the phylogenetic structure of the bats (Teeling et al., 2005). This may have developed as a mechanism of avoiding direct competition between species. This would be particularly key where species are roosting together, and potentially exploiting resources within the same foraging areas. In these data, the morphologically cryptic species appear to *P. pygmaeus* and *P. pipistrellus* cluster into different dietary guilds. However, the Pianka's overlap between the two species is 0.886, indicating that there is still large potential for competition between the species; >60% is considered "biologically significant" (Bethea et al., 2006).

M. nattereri has the highest level of overlap with other Great British bat species, with an average Pianka's overlap of 0.878. This suggests that this bat species appears to be potentially the most at risk of competition from other species.

4.5.3. Seasonal variations

Seasonal variation in Arthropod availability is reflected in the bat diets; figure 4.14 indicates that the proportions of Lepidoptera may be higher in summer and autumn than in spring and winter. This may reflect the increase in lepidopteran activity in summer and when the weather is warm (Jonason et al., 2014). It is thought that climate change will increase Great British average temperatures, which would drive an increase in Lepidopteran activity and migration (Cannon, 1998, Sparks et al., 2005, Sparks et al., 2007, Turner et al., 1987). Opportunistically feeding bats may increase the proportions of Lepidoptera consumed as a result of climate change. Conversely, the proportions of Diptera appear to be in the samples submitted in winter. This may be as a result of the lack of Lepidoptera availability in winter, as Lepidoptera have (on average) larger body sizes, so may be preferentially selected in summer. However as previously discussed, further faecal samples would be needed to confirm this.

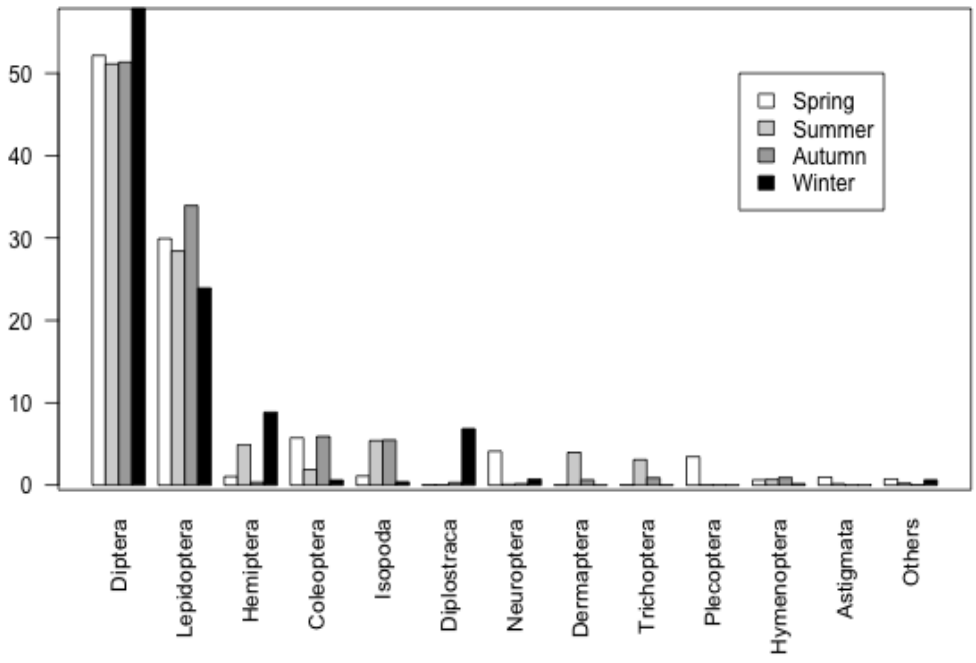


Figure 4.14: The proportions of Arthropoda reads of the seasons at order level

Grouped by when samples were submitted to the BatID service, into spring, summer, autumn, and winter. Code adapted from appendix D.2.10. Spring- n=23, Summer- n=30, Autumn- n=33, Winter- n=17.

Seasonal variation may be particularly important in the diets of generalist feeders that are opportunist in their feeding.

4.5.4. Variations across Great Britain

Figure 4.14 shows the Arthropoda representation in the samples across Great Britain. Most strikingly, is the very high proportion of Diptera in the Scottish samples, compared to the very low proportions of Lepidoptera. The species of bats with ranges extending to Scotland include *M. daubentonii*, *M. mystacinus*, *M. nattereri*, *N. leisleri*, *N. noctula, P. nathusii*, *P. pipistrellus*, *P. pygmaeus*, and *P. auritus*. These fall into the two major guilds, suggesting that the location is more important than the species in determining the ratio of Diptera and Lepidoptera in Scotland.

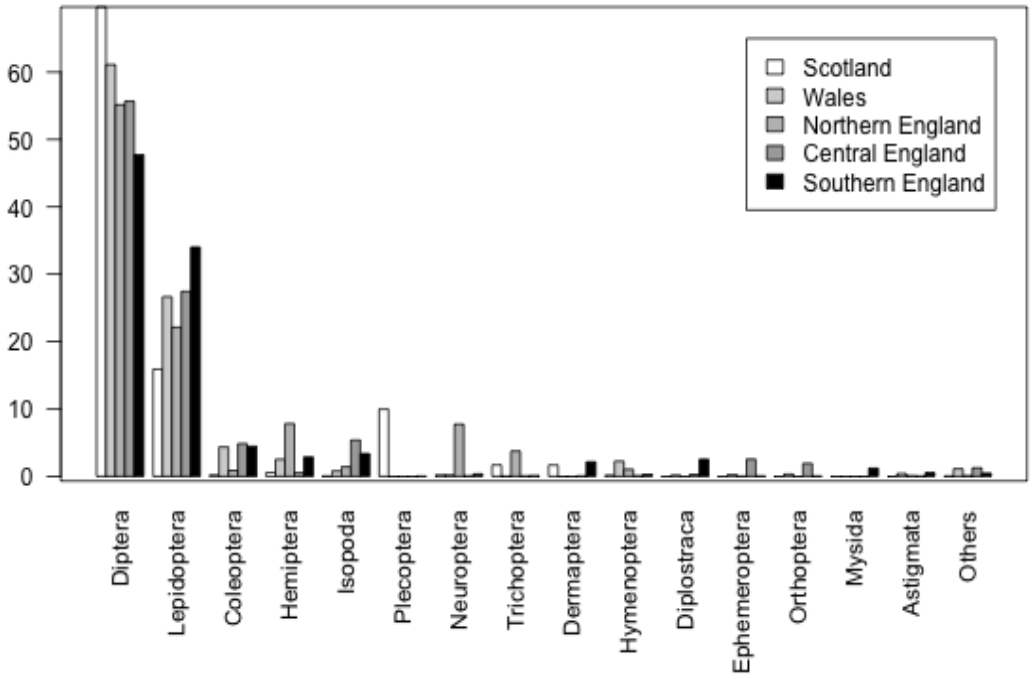


Figure 4.15: The proportions of Arthropoda reads of across Great Britain
Grouped into Scotland, Wales, Northern England, Central England, southern England. Code adapted from appendix D.2.10. Spring- n=23, Summer- n=30, Autumn- n=33, Winter- n=17.

Lepidoptera and Diptera are clearly vital in the diets of all bat species. This is reflected in figure 4.16, where Lepidoptera and Diptera explain most of the variation in principal components 1 and 2.

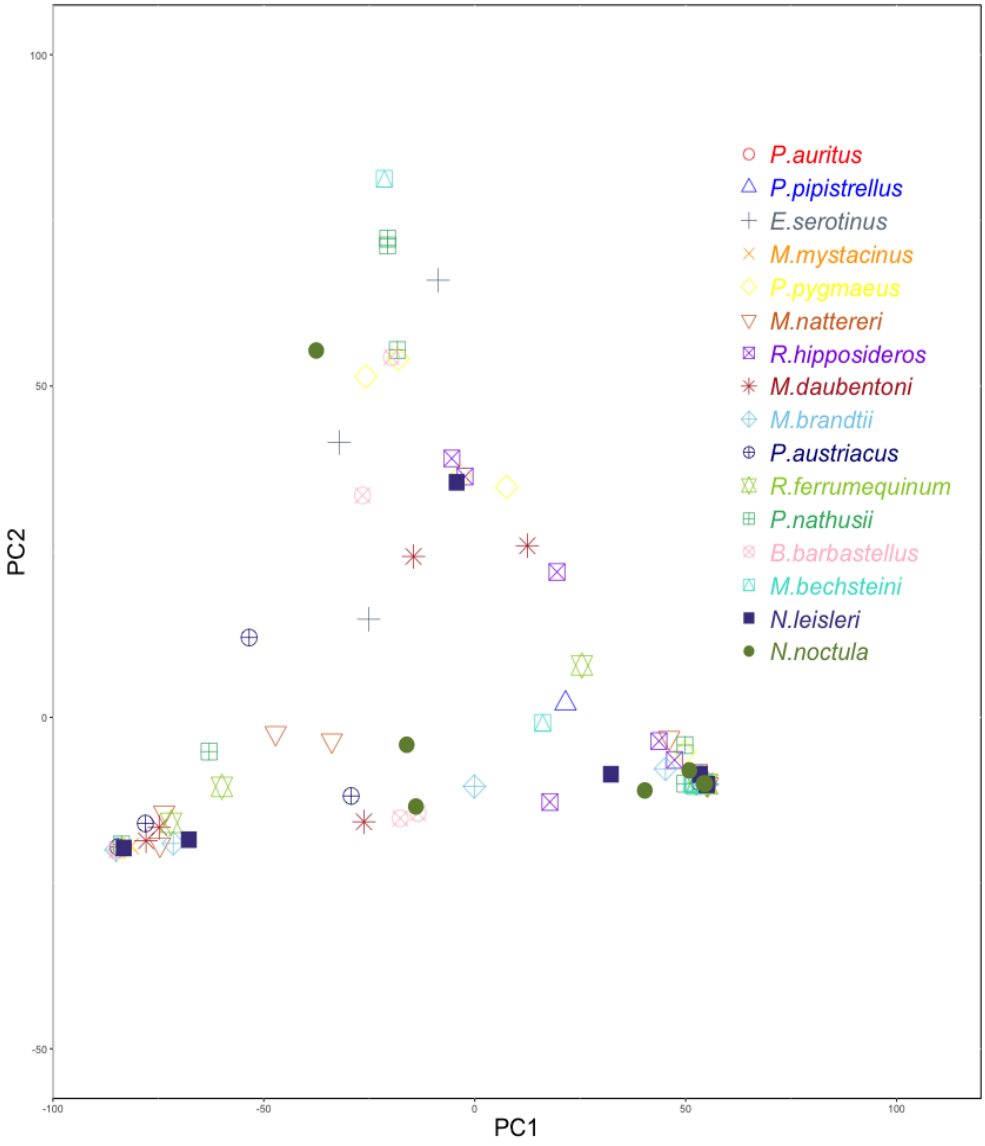


Figure 4.16: Distribution of the Great British bat diets in principal component space

Derived from the proportions of diet species within the libraries. Principal components one and two are shown; together they explain 77.43% of the variation (PC1- 51.68%, PC2- 25.75%). Component one is defined primarily by Lepidoptera (loading of -0.661) and Diptera (loading of 0.750), and component two by Isopoda (0.266), Lepidoptera (0.662), Coleoptera (0.356), Diptera (0.563), Hemiptera (-0.124). A summary and loadings can be found in appendix E.4.1. Figure code adapted from D.2.11.

4.5.5. Future work

It is unlikely that any primer set designed will be able to capture the full breadth of the Arthropoda, as the majority of Arthropoda CO1 genes have

not been sequenced. As of April 2016 103,503 species were represented on the BOLD database (Ratnasingham and Hebert, 2007), whereas there are ~6.8 million species (range 5.9-7.8 million) predicted terrestrial Arthropoda species globally (Stork et al., 2015). As sequence databases are continually updated, it is advisable that primers used for metabarcoding are regularly assessed and, if necessary, re-designed.

In this dataset, there were only 72 sequences matched primer >18_F_Arthropoda_COI, suggesting that it is not useful, and should not be used in future. Additionally, primer >11_R_Arthropoda_COI amplified high proportions (90.16%) of spurious sequences, and should also be excluded from future work.

4.6. Conclusions

The intra-specific variation is large (figure 4.16), as has been previously seen (Boyles and Storm, 2007). PIA is useful for metabarcoding data, but it appears to be overly conservative with some data. It is more appropriate, and less computationally heavy, to use PIA on metagenome data. The use of PIA to validate data will be further explored in chapter five.

The Arthropoda data from this chapter will be compared to the data collected from the literature (chapter two) and from the shotgun metagenome (chapter three) in chapter five and chapter seven will place these data in an ecological context.

Chapter 5 : Synthesis of the different datasets

5.1. Abstract

There are a variety of methods available for studying bat diets. In this chapter, all three diet datasets compiled in this thesis are compared; the data compiled from the meta-analysis of the literature, from the shotgun metagenome data, and from the metabarcoding data.

The full diet literature analysis identified 18 orders of diet species, of which 2 were novel (not seen in either of the other datasets). The metagenome data identified 19 orders, of which 1 was novel. The metabarcoding data identified 38 orders, and the highest number of novel orders (17).

Metabarcoding data is, at present, the most useful method of identifying bat diets from guano. However, as sequencing power increases (and costs decrease), and the reference sequence databases are expanded, this is likely to change.

5.2. Introduction

The main aim of this chapter is to compare the two methods used in this thesis: shotgun metagenomics and metabarcoding. These data are also compared against the data reported in the literature that pertains to the diets of the bats studied (chapter 2).

As discussed in section 1.4, there are advantages and disadvantages in the use of both methods. In this chapter, we directly compare the results gathered by both methods, before and after Phylogenetic Intersection Analysis (section 3.3.2.5.1).

5.3. Materials and methods

The analysis undertaken in this chapter was as described in 3.3.2 and 3.3.3. Mantel tests were completed using the mantel.rtest function from

the R ade4 package (Dray and Dufour, 2007). All tests were based on 9,999 permutations.

5.4. Results

5.4.1. The diet data returned

The metagenome dataset returned 6,411,192 reads, of which 2,656,124 were assigned during BLAST analysis, with 26,569 assigned to Arthropoda.

The metabarcoding forward read dataset returned 10,849,278 reads, of which 10,779,663 were assigned during BLAST analysis, with 9,514,856 were assigned to Arthropoda.

Table 5.1: Number of Arthropoda assignments of the three datasets before PIA
From the literature review, from the metagenome, and from the metabarcoding datasets after primer trimming. Showing order, family, genus, and species.

Taxa	Literature review	Shotgun metagenome	Metabarcoding
Order	18	19	38
Family	N/A	79	264
Genus	N/A	107	538
Species	N/A	114	634

5.4.2. Novel taxa not previously identified in bat diets

In all three datasets there were 39 orders identified in bat diets. How well these orders are represented in the NCBI dataset is shown on table 5.1. The average representation on the NCBI databases for the orders identified is 465,152, with the average number of COI records for each order being 31,496. Where an order has a lower number of records on the NCBI databases, the likelihood of molecular methods identifying these orders is lower.

Table 5.2: The orders identified in each bat species diet before PIA.

Also showing the representation on the NCBI nt database. Showing all of the results from the NCBI database (column two) and the number of records assigned to the cytochrome c oxidase gene (column three). ✓ - <1% of the whole dataset, ✓✓ - 1-<10%, ✓✓✓ - 10-100%.

Order	NCBI nt	NCBI nt COI	Literature	Metagenomics	Metabarcoding
Amphipoda	116,899	11,074			✓
Araneae	745,838	42,588		✓	✓
Archaeognatha	1,562	164			✓
Astigmata	10,559	1,389		✓✓	✓
Blattodea	30,021	899	✓		
Calanoida	534,935	7,298			✓
Chordeumatida	154	124			✓
Coleoptera	925,808	106,642	✓✓✓	✓✓	✓✓
Collembola	99,382	24,486			✓
Decapoda	1,008,125	33,442		✓	✓
Dermaptera	27,403	506	✓✓	✓	✓✓
Diplostraca	573,075	5,240			✓✓
Diptera	3,417,228	432,091	✓✓✓	✓✓✓	✓✓✓
Ephemeroptera	231,47	12,275	✓	✓	✓
Euphausiacea	3,215	1,611			✓
Harpacticoida	76,668	455		✓✓	✓
Hemiptera	1,727,688	57,774	✓✓	✓✓	✓✓
Hymenoptera	5,925,649	143,480	✓✓	✓✓	✓
Isopoda	33,121	3,706	✓		✓✓
Ixodida	984,831	1,467		✓	
Julida	397	187			✓
Lepidoptera	2,317,997	332,980	✓✓✓	✓✓✓	✓✓✓
Lithobiomorpha	8,097	151			✓
Mantodea	4,752	375			✓
Megaloptera	72,092	560			✓
Mesostigmata	123,258	7,739		✓✓	✓
Mysida	1,889	1,014			✓
Neuroptera	72,841	2,521	✓✓	✓	✓✓
Odonata	43,643	6,323			✓
Opiliones	10,049	2,473	✓	✓	
Oribatida	11,510	10,075			✓
Orthoptera	176,144	9,089	✓	✓	✓
Pantopoda	2,261	947		✓	✓
Pedunculata	3,321	1,261			✓

Plecoptera	27,600	6,393	✓		✓
Pseudoscorpiones	835	299	✓		
Psocoptera	10,228	7,880	✓		✓
Scorpiones	17,340	1,579			✓
Sessilia	15,738	9,083		✓	✓
Siphonaptera	93,552	336	✓		✓
Thysanoptera	56,905	9,542			✓
Trichoptera	200,632	25,330	✓✓	✓	✓

In the literature data, there were two Arthropoda orders identified that were not present in either of the other datasets; the Blattodea (which contains Termites and Cockroaches) and the Pseudoscorpiones (the false or book scorpions). Whilst there are three species of Blattodea native to Great Britain (Barnard, 1999) along with a number of invasive species (Alexander et al., 1991), there were no Blattodea identified in the literature data collected in Great Britain. The Blattodea were identified in samples collected in Germany from *M. bechsteinii*, *M. nattereri*, and *P. auritus* (Roswag et al., 2015). These three species have late emergence times compared to the other species studied, except for *M. daubentonii* (see tables 1.1 and 5.2) (Jones and Rydell, 1994). This is consistent with the nocturnal activity of the Blattodea (Rust and Reierson, 2007). There was only one reported case of Pseudoscorpiones in in a bat diet, in a *M. daubentonii* from Ireland (Flavin et al., 2001). Pseudoscorpiones was a very small proportion of the reported diet. These orders have much lower than average representation on the NCBI nt database (table 5.2), which may cause them to have been absent from the molecular data.

The metabarcoding dataset returned the highest number of novel Arthropoda orders (17). These include the Amphipoda, Archaeognatha, Calanoida, Chordeumatida, Collembola, Diplostraca, Euphausiacea, Julida, Lithobiomorpha, Mantodea, Megaloptera, Mysida, Odonata, Oribatida, However, none of these were highly represented in the metabarcoding dataset; each was <1% of the whole metabarcoding dataset, except Diplostracia, which comprised 1.3% of the metabarcode dataset (see section 5.4.3 and table 5.3).

In the metagenome dataset there was one novel order: the Ixodida, which comprises 0.782% of the metagenome dataset. The lack of metabarcoding data may be due to the high number of Ixodida hits on the NCBI that are not assigned to COI (table 5.1). For orders with a high ratio of total NCBI data compared to NCBI COI data (i.e. those orders which have species which genomes sequenced), metagenomic approaches may be more sensitive than metabarcoding approaches.

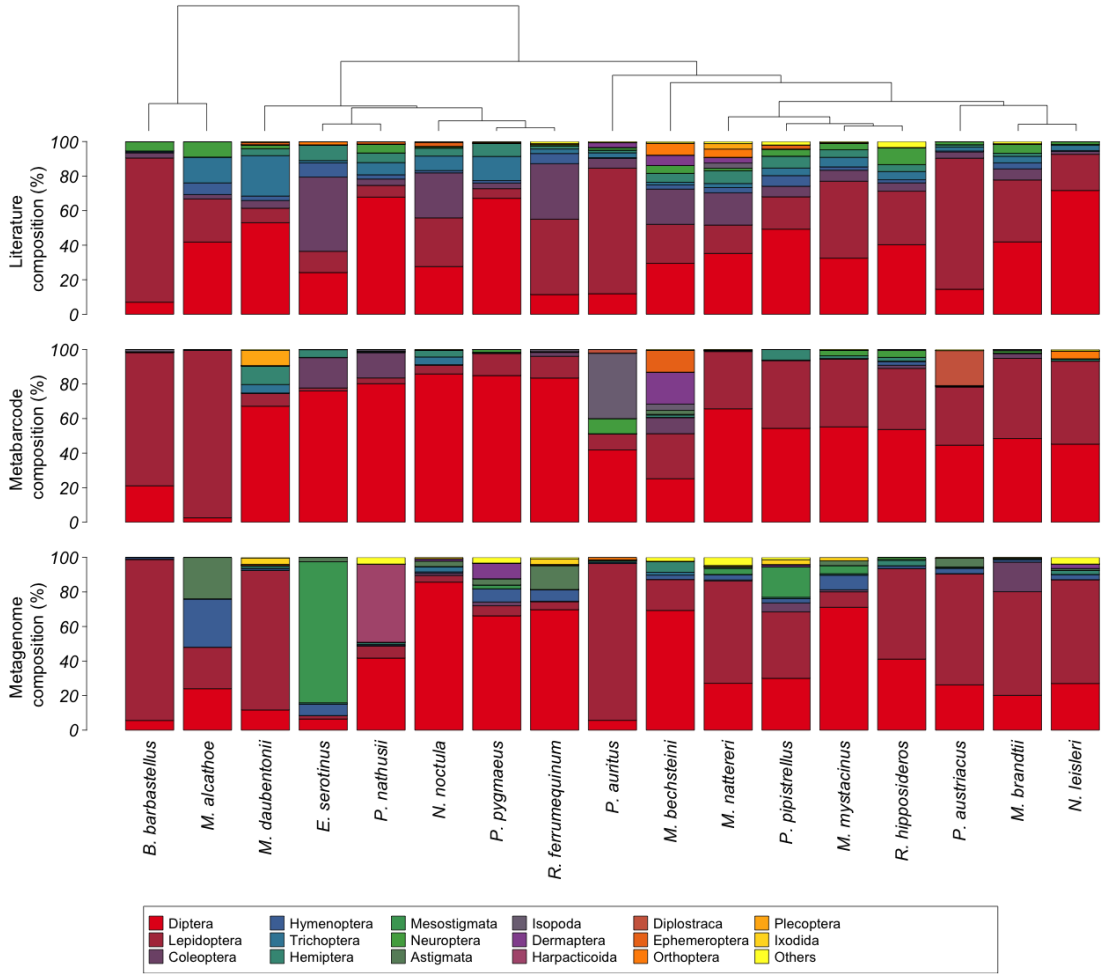


Figure 5.1: The diet of each bat species at order level

As determined through a. the literature review, b. by metabarcoding, and c. by metagenome analyses before PIA. Prey taxa have been grouped by order. Diets sorted by the metabarcoding data using complete-linkage clustering using the R hclust package. Code adapted from appendix D.2.2.

In the literature data, the highest represented order was the Diptera, with 37.87% of the orders represented in the dataset, with Lepidoptera comprising 32.47%. Diptera comprised between 7.01% (*B. barbastellus*) and 71.69% (*N. leisleri*) of the species datasets (figure 5.1) from the literature (section 2.4.2). Lepidoptera comprised between 5.59% (*P. pygmaeus*) and 83.46% (*B. barbastellus*). In the metabarcode data the highest represented order was also Diptera with 55.02% of the dataset. Lepidoptera comprised 31.04% of the metabarcode dataset. Diptera comprised between 2.54% (*M. alcathoe*) and 85.83% (*N. noctula*), and Lepidoptera between 1.70% (*E. serotinus*) and 96.95 (*M. alcathoe*) (section 4.4.5). Finally, Lepidoptera comprised 39.61% of the metagenome dataset, with Diptera comprising 36.99%. The Lepidoptera comprised between 4.61% (*R. ferrumequinum*) and 93.12% (*B. barbastellus*), and the Diptera between 5.64% (*B. barbastellus*) and 85.70% (*N. noctula*) in the metagenome dataset (section 3.4.7).

B. barbastellus constantly has the highest proportion of Lepidoptera in its diet (figure 5.1) in the literature, metabarcode and metagenome datasets. As a result, it has a low niche breadth and dietary diversity compared to the other species (figure 5.2).

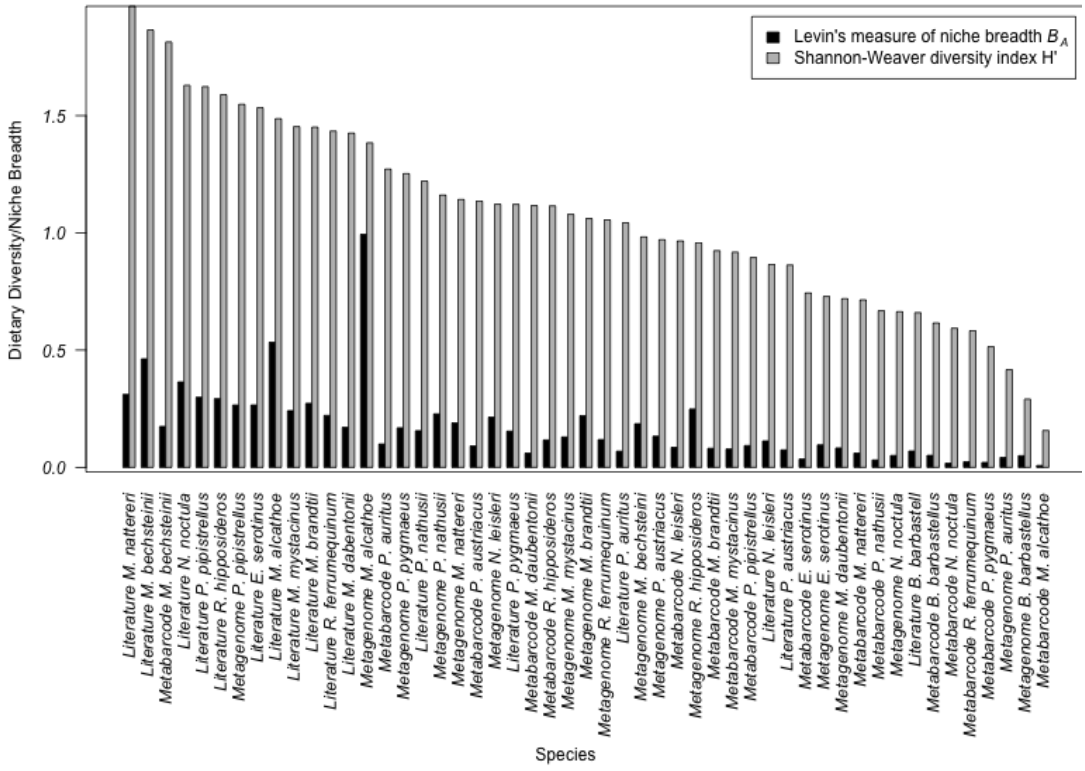


Figure 5.2: Dietary diversity and niche breadth of each bat species before PIA
The dietary diversity calculated using Shannon-Weaver diversity index (H) (grey) and niche breadth calculated using Levins standardised index (B_A) (black). Arthropod data assigned at order level. Code adapted from appendix D.2.3.

The literature data returns the highest average dietary diversity (Shannon-Weaver diversity index (Shannon and Weaver, 1948)) for all of the species at 1.36, the metagenome dataset has a Shannon-Weaver index of 0.97, and metabarcoding has an index of 0.86. The literature dataset also has the highest average niche breadth (Levin's index (Levins, 1968)) of 0.24%, then the metabarcode dataset at 0.2, then the metabarcoding dataset at 0.07. *M. alcathoe* has a large disparity between the niche breadth and dietary diversity, which is likely due to the low quantity of data assigned to this species, particularly in the metagenome dataset. If *M. alcathoe* is included in the dataset, the multiple R^2 for Levins values against Shannon-Weaver values was 0.4102, whereas if *M. alcathoe* was excluded, the value increases to 0.7199.

Lepidoptera and Diptera form the basis of the diets of the Great British bats studied (figure 5.1). This allows the formation of three dietary guilds

(figure 5.3). The first (figure 5.3, top right) has higher proportions of Lepidoptera. This guild contains all three *B. barbastellus*, the *P. auritus* and *P. austriacus* literature and metagenome datasets. This guild also contains the metabarcoding data for *M. alcathoe*, the metagenome data for *M. nattereri*, *N. leisleri*, *M. brandtii*, and *M. daubentonii.*

The second guild (figure 5.3, bottom left) contains higher proportions of Diptera and lower proportions of Lepidoptera. This guild contains all three *P. pygmaeus* datasets, two out of three datasets for *M. daubentonii*, *R. ferrumequinum*, *P. nathusii*, *N. noctula*, and *M. nattereri*. Between these guilds, there is a mixed guild, which includes the diets that have a dependency on both Lepidoptera and Diptera, which contains a large number of bat datasets.

The metagenome *M. alcathoe* and *E. serotinus*, and literature *E. serotinus* data do not cluster well into the other guilds. In the case of *M. alcathoe* this is likely due to the proportionally lower amounts of data for *M. alcathoe*. In the case of *E. serotinus* the high proportions of Mesostigmata in the metagenome data sets it apart from the other species. One reason that Mesostigmata were not seen in the metabarcode dataset is that the relative proportion of Mesostigmata COI entries on the N.C.B.I database compared to the rest of the Mesostigmata is only 6.28%. This is compared to an average proportion of 20.07% Arthropoda COI entries out of the total Arthropoda entries (table 5.2). An alternative reason for the lower representation of the Mesostigmata in the *E. serotinus* metabarcode dataset is that it is possible that the metabarcoding primers are not as good at amplifying Mesostigmata DNA as other orders. One final reason is that in the Mesostigmata data in the *E. serotinus* metagenome dataset may be false assignment, however, after PIA (section 5.4.3) these data were retained, so this is unlikely in this case.

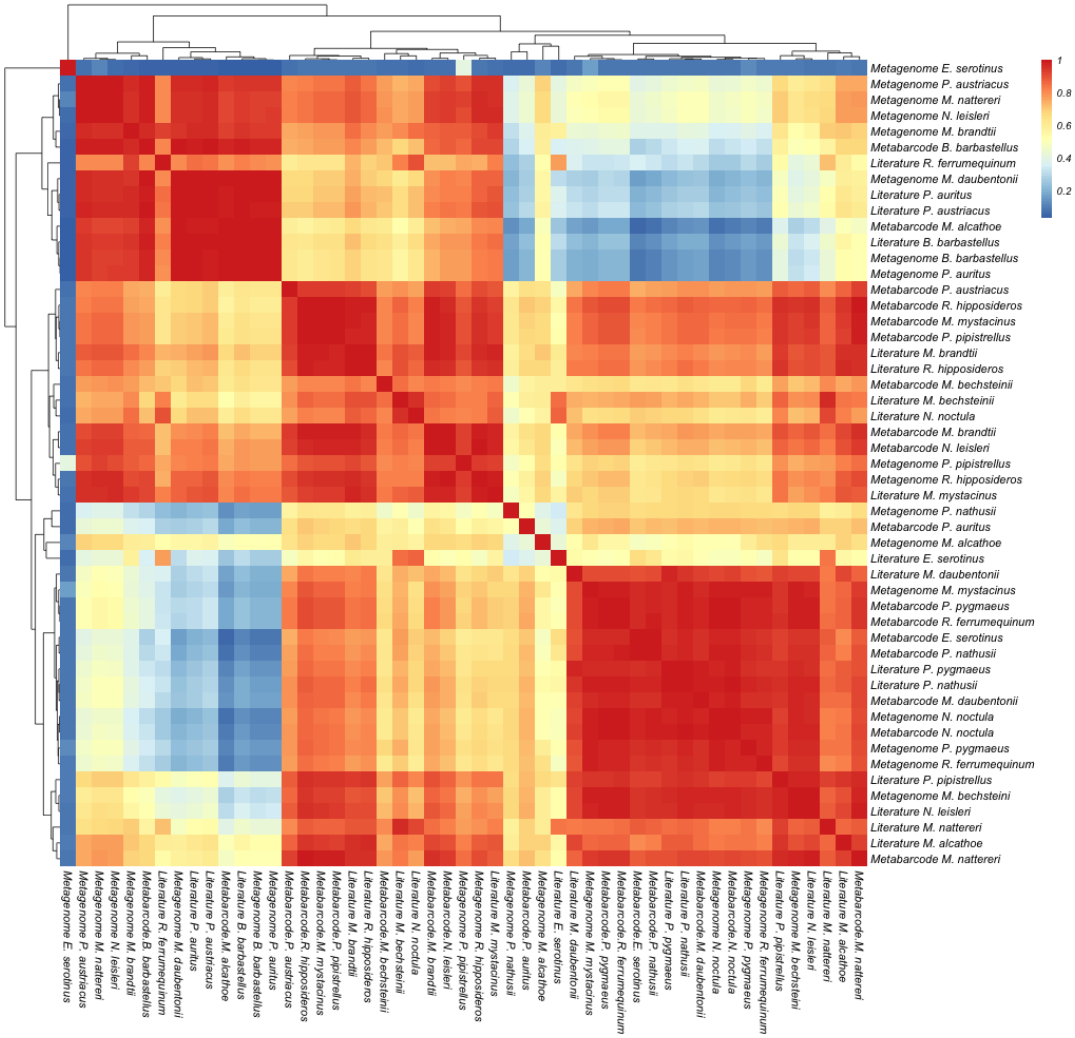


Figure 5.3: The dietary overlap of each bat species

Using Arthropoda data at order level before PIA. Calculated using Pianka's index of niche overlap. A value of 1 (red) suggests that the diets are identical, whilst a value of 0 (blue) indicates that there is no overlap. Species have been hierarchically clustered by niche overlap similarity. Code adapted from appendix D.2.4.

The mean overlap was 0.683, with the highest overlap between different species datasets (0.999) being between *P. pygmaeus* metabarcoding data and *R. hipposideros* metabarcoding data, and the lowest overlap (0.0261) between the *M. alcathoe* metabarcode data and the *E. serotinus* metagenome data.

5.4.3. The impact of phylogenetic intersection analysis

After PIA, 70.4% of the metagenome Arthropoda data was discarded, compared to the 33.57% discarded from the metabarcode Arthropoda data. The lower proportion of the metabarcode data that was discarded reflects the better representation of barcoding targets on the NCBI databases. Some of the metabarcode data that was discarded was due to the lack of diversity within the metabarcoding datasets. Additionally, the lack of diversity within the metabarcoding targets resulted in a much longer running time when running metabarcoding data compared to when running shotgun metagenome data.

In the metabarcode dataset, before PIA there were 38 orders represented (table 5.1), whereas after PIA there were 24 orders represented. Conversely, in the metagenome dataset, before PIA there were 19 orders represented, whereas after PIA there were 26 orders represented. This suggests that in metagenome datasets, the rate of false assignment (before PIA) is high, with the draw of highly represented (i.e. model) organisms being very powerful. When working with metabarcode data, PIA tends to discard some orders, particularly where diversity is low.

After PIA, a number of orders had increased representation in both datasets (table 5.3): in the Araneae, Blattodea, Coleoptera, Cryptostigmata, Isopoda, Neuroptera, and the Orthoptera. The orders which were decreased, or removed in both datasets were the Dermaptera, Harpacticoida, Lepidoptera, and Pantopoda.

Table 5.3: The changes in orders identified in the full metagenome and metabarcoding datasets after PIA

↑– Increased after PIA, ↓- decreased after PIA, X -removed after PIA, ✔-added after PIA.

Order	Metagenome data after PIA	Metabarcode data after PIA
Amphipoda		X
Araneae	↑	↑
Archaeognatha		X
Astigmata	↑	↓

Blattodea	✔	✔
Calanoida		X
Chordeumatida		X
Coleoptera	↑	↑
Collembola		X
Cryptostigmata	✔	✔
Decapoda	↓	↑
Dermaptera	↓	↓
Diplostraca		X
Diptera	↑	↓
Ephemeroptera	↑	↓
Euphausiacea		X
Harpacticoida	↓	X
Hemiptera	↓	↑
Hymenoptera	↓	↑
Isopoda	✔	↑
Isoptera		✔
Ixodida	↓	
Julida		X
Lepidoptera	↓	↓
Lithobiida		✔
Lithobiomorpha		X
Mantodea	✔	X
Megaloptera		↑
Mesostigmata	↑	X
Mysida		X
Neuroptera	↑	↑
Odonata		↑
Opiliones	↑	
Oribatida		X
Orthoptera	↑	↑
Pantopoda	↓	X
Pedunculata		X
Plecoptera	✔	↓
Psocoptera	✔	X
Raphidiodea	✔	
Scorpiones	✔	X
Sessilia	X	↑
Siphonaptera		X
Thysanoptera		X
Trichoptera	↓	↑

Before PIA, there were a number of marine and freshwater associated arthropods in both datasets, which are unlikely to be true diet species. After PIA, the majority of these were removed (the Amphipoda, Calanoida, Diplostracia, Euphausiacea, Mysida, and Pendunculata). The Harpacticoida and Pantopoda were significantly reduced from the metagenome dataset (both were removed from the metabarcode dataset). There was a slight increase in the Sessilia (barnacles) in the metabarcode dataset, from 0.0037% to 0.0042% (although this is still very low). Decapoda were increased in the metabarcode dataset (0.022% to 0.31%) and reduced in the metagenome dataset (0.13% to 0.0.22%), which may be as a result of their high representation on the NCBI database (table 5.2).

The increase in Decapoda data may be a misassignment of other data, due to the lower diversity in the metabarcode dataset than in metagenome data. Overall, 76.83% of the aquatic Arthropoda data was removed from the metabarcode dataset, and 98.59% from the metagenome dataset. This confirms that the marine arthropod data before PIA was as a result of the misassignment of data where incomplete databases are used, and validates the use of PIA, particularly for metagenome data (see section 4.4.6.1). Alternatively, it is possible that this data is correctly assigned and, whilst these orders may not have been preyed upon directly by the bats, it is possible that they were present in the diets of predatory arthropods that had been eaten by the bats.

Table 5.4. shows the taxonomic levels that the Arthropoda data are assigned to for both methods before and after PIA. Before PIA, a large amount of the data from both datasets is assigned at species level. This is as a result of the tendency of methods such as BLAST to assign data to terminal leaf nodes, and results in misassignment of data to over represented species on the database used (section 3.3.2.5.1). After PIA, the data are shifted to be assigned at higher taxonomic levels. PIA does not assign data to terminal leaf nodes, however small proportions of the

data are assigned at species level. This is as a result of data being initially assigned to subspecies. PIA rarely assigns data at levels higher than Class as the genetic distance between species belonging to different Phyla etc., is such that intersections are not found between such species.

Table 5.4: Taxonomic level of data assignments for metagenome and metabarcode Arthropoda data before and after PIA

Taxonomic level	% metagenome data before PIA	% metagenome data after PIA	% metabarcode data before PIA	% metabarcode data after PIA
Superkingdom	0.69	0.00	0.00	0.00
Phylum	3.31	0.00	1.53	0.00
Class	8.94	58.87	8.04	53.32
Order	21.67	14.55	36.43	15.60
Family	15.43	21.47	21.54	16.84
Genus	6.38	4.60	2.46	13.83
Species	43.58	0.51	29.99	0.41

After PIA, the metabarcode data has data assigned lower down the taxonomic tree (closer to the terminal leaf) than in the metagenome data. This is due to the greater completeness of the metabarcode reference datasets.

There is a large disparity between the results for the metabarcode and metagenome data after PIA (figure 5.4). This is likely due to the distinct difference in the amount of data for each dataset: the number of Arthropoda hits in the metagenome dataset is far lower than in the metabarcoding data.

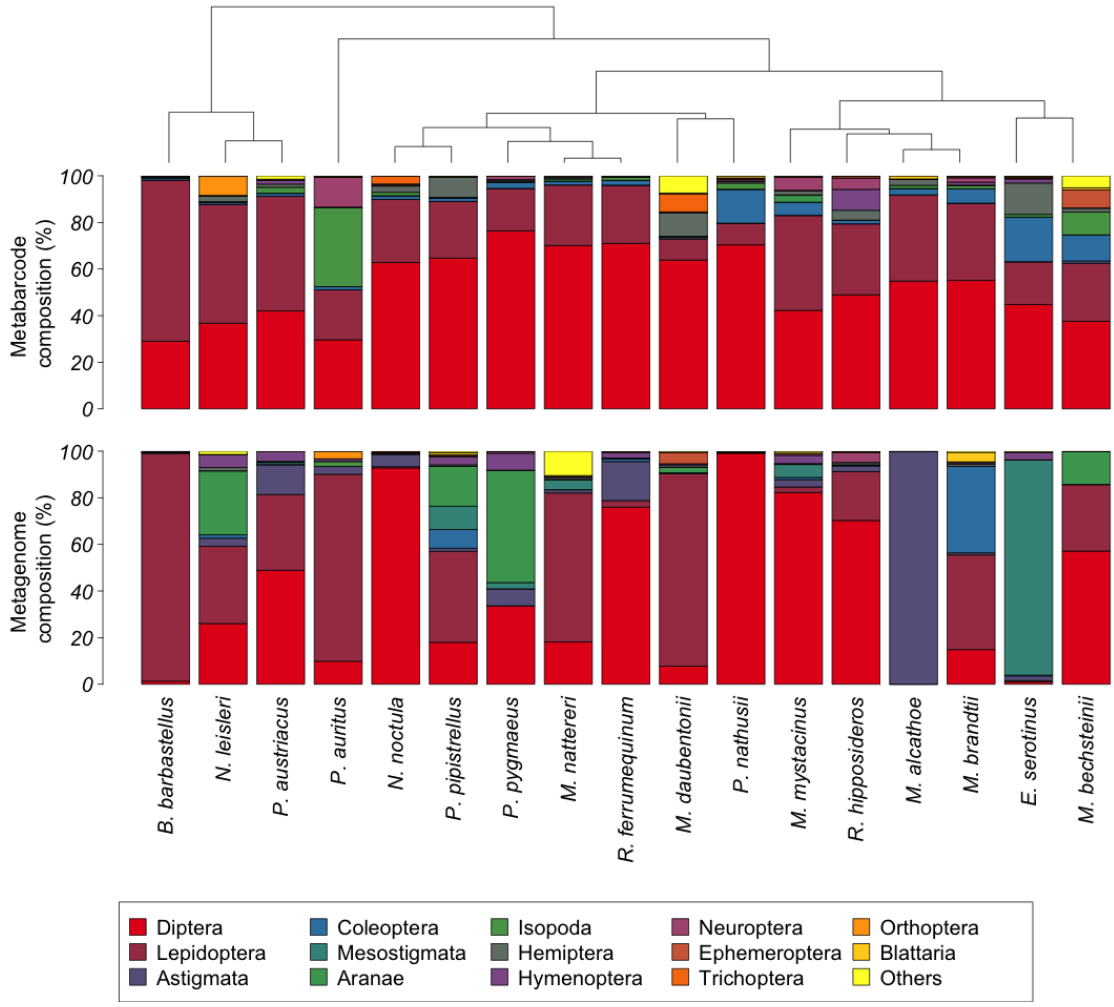


Figure 5.4: The diet of each bat species at order level

As determined through a. by metabarcoding, and b. by metagenome analyses after PIA. Prey taxa have been grouped by order. Diets sorted by the metabarcoding data using complete-linkage clustering using the R hclust package. Code adapted from appendix D.2.2. Others includes Cryptostigmata, Decapoda, Dermaptera, Harpacticoida, Isoptera, Ixodida, Lithobiida, Mantodea, Megaloptera, Odonata, Opillones, Pantopoda, Plecoptera, Psocoptera, Raphidiodea, Scorpiones, Sessilia.

The average dietary diversity of the metabarcode data was 2.48 (figure 5.5) and the average niche breadth was 1.11. The species with the highest dietary diversity was *M. bechsteinii* and the species with the highest niche breath was *P. auritus.* The lowest dietary diversity was *B. barbastellus,* and the lowest niche breadth was *P. pygmaeus*. In the metagenome dataset, the average dietary diversity was 2.04 and the average niche breadth was 0.83. The species with the highest dietary diversity *P. pipistrellus,*

the species with the highest niche breadth was *M. bechsteinii*. The species with the lowest dietary diversity and breadth was also *M. alcathoe,* followed by *P. nathusii*.

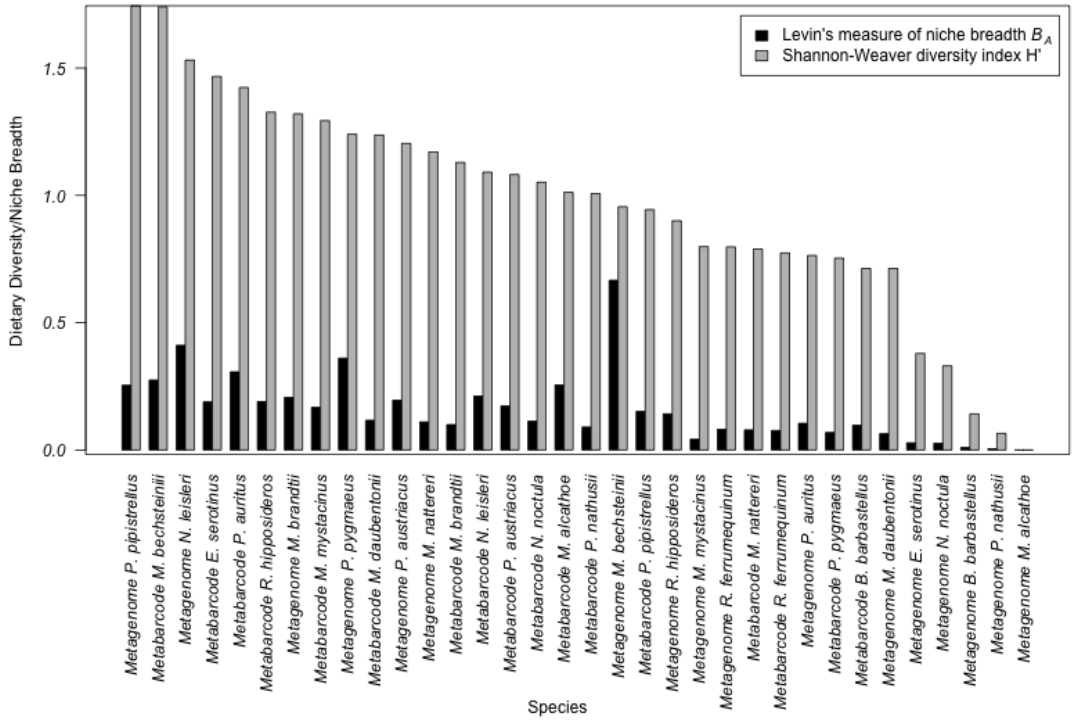


Figure 5.5: Dietary diversity and niche breadth of each bat species after PIA
The dietary diversity calculated using Shannon-Weaver diversity index (H) (grey) and niche breadth calculated using Levins standardised index (B_A) (black). Arthropod data assigned at order level. Code adapted from appendix D.2.3.

M. alcathoe may be skewing the metagenome trends due to the lack of data. When *M. alcathoe* is excluded from the datasets, the average dietary diversity of the metabarcode data was 0.15 and the average niche breadth was 1.11. In the metagenome dataset, the dietary diversity without *M. alcathoe* was 0.87, and the niche breadth was 0.17.

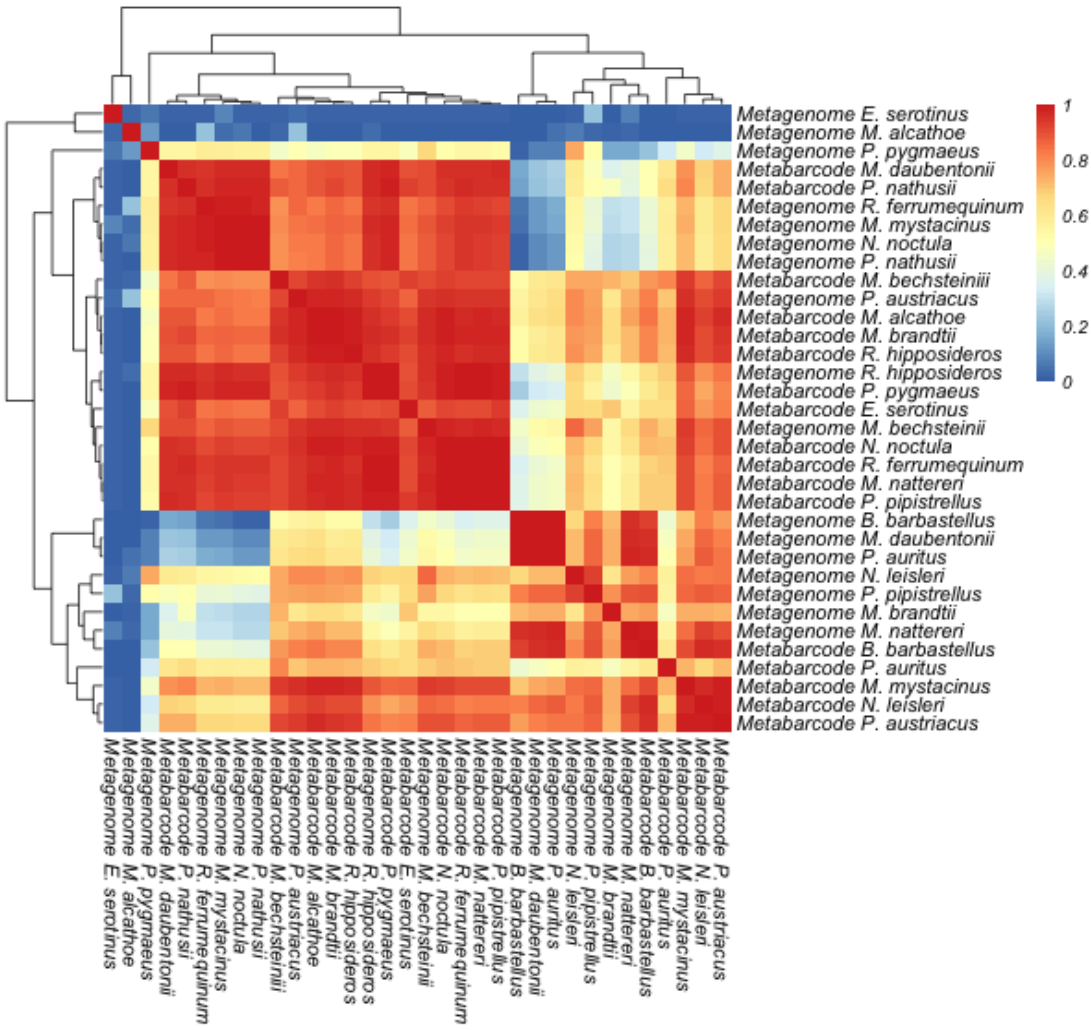


Figure 5.6: The dietary overlap of each bat species with Arthropoda at order level after PIA

Calculated using Pianka's index of niche overlap. A value of 1 (red) suggests that the diets are identical, whilst a value of 0 (blue) indicates that there is no overlap. Species have been hierarchically clustered by niche overlap similarity. Code adapted from appendix D.2.4.

After PIA, the guild structure (as defined using Pianka's index) in figure 5.6, shows similar trends to before PIA (figure 5.3). There are still three guilds: one based on Diptera (top left group), one based on Lepidoptera (bottom right) and one with a balance of Diptera and Lepidoptera (middle group). Again, the metagenome *E. serotinus* and *M. alcathoe* datasets do not cluster well with the other groups.

5.5. Discussion

Metabarcoding approaches return far more Arthropoda data than shotgun metagenome approaches. The two datasets show positive Mantel associations when comparing the two datasets for each species (metagenome vs metabarcode) (r= 0.89, *p*=8e-04, after 9,999 permutations) (Mantel, 1967). This falls slightly after PIA (r=0.85, *p*=0.8e-04, after 9,999 permutations). This may be due to the disparities between the coverage of whole metagenome and metabarcoding targets in the databases available (table 5.2). I would suggest that this highlights the need for caution when using any dataset where incomplete reference datasets are used. Due to the larger volume of metabarcode data, and the more complete reference datasets, it is likely that metabarcode data is, at present, more representative of the true bat diets.

5.5.1. Assessing the use of PIA on metabarcoding and metagenome datasets

The increased reliability of the metabarcode data in comparison to the metagenome dataset is further demonstrated through the use of Phylogenetic Intersection Analysis. Mantel tests on the full metabarcode dataset before and after PIA gives a strong observed correlation, suggesting robust support for the metabarcoding data through PIA (r= 0.99, *p*=1e-04, 9,999 permutations). Comparatively, the support for the full metagenome data by the PIA is lower (although still significant) (r=0.202, *p*= 1e-04, 9,999 permutations). The lower proportion (33.57%) of barcoding data discarded after PIA compared to the 70.4% of metagenome data discarded, further demonstrates the importance of using complete reference datasets. As reference databases are added to, more metagenome data will be retained in analyses, increasing the efficacy of shotgun metagenome approaches in dietary analyses. This is particularly relevant as metagenome approaches have the potential to be more sensitive than metabarcoding approaches as it exploits more of the DNA laid down by the organisms (section 3.4.8 and 3.3.2.5.2). This is highlighted in table 5.5, which shows that despite the larger volumes of Arthropoda data returned by the metabarcode data, the metagenome

approach is more efficient at identifying distinct Arthropoda orders. This is despite the far higher proportion of data discarded during metagenome approach (section 5.4.1). This is also reflected after PIA.

Table 5.5: The numbers of Arthropoda orders identified by each method per 1,000 non-degenerate Arthropoda reads

Before and after PIA.

	Metabarcode	Metagenome
Non-degenerate Arthropoda reads (NDR)	650,400	26,968
Orders before PIA	38	19
Orders/1,000 NDR before PIA	0.058	0.7045
Orders after PIA	24	26
Orders/1,000 NDR after PIA	0.036	0.964

5.5.2. Future work

In both the metabarcoding and metagenome datasets, *M. alcathoe* is under represented. In future work, additional *M. alcathoe* samples should be sourced. These were not available in this study due to the scarcity of this species (Jan et al., 2010).

5.6. Conclusions

Metabarcoding data is, at present, the most useful method of identifying bat diets from guano. However, as sequencing power increase (and costs decrease), and the reference sequence databases are expanded, this is likely to change. When using any method for analysing diets, the use of technical replicates is crucial.

Chapter six explores the use of guano morphology to identify bat species, and chapter seven analyses the impact of diet on guano morphology. Chapter seven will then sum up the findings of this thesis, and place them

in the context of our knowledge about the individual bat species, their behaviours, and habitats.

Chapter 6 : A critical appraisal on the use of guano morphology to identify species

6.1. Abstract

Accurate species identification is a crucial part of studying and data-driven conservation of bats. There are a number of widely used methods in the study of bats, each with their own particular pitfalls and problems. We tested whether it was possible to distinguish between Great British bat species on the basis of guano morphology alone.

Guano from 16 of the Great British bat species was collected and subjected to mitochondrial DNA analysis order to confirm the species identity, after which, various physical parameters were measured and subjected to principal components analysis.

We show that measuring length, diameter, colour, particle size and nodulation of guanos provide the best resolution for differentiating between species. We further demonstrate a clear shift between guano parameter ranges in dry and fresh samples. With species traditionally associated with distinctive guano morphology (such as *E. serotinus*), we saw areas of the guano ranges where there was a possibility of incorrectly identifying the species. Indeed, there was no single species that could be unequivocally identified for all areas of its range. Whilst for some species, using guano morphology to identify a sample may be useful; in other cases, the chance of misidentifying a sample may be prohibitively high.

6.2. Introduction

Traditionally, the monitoring and recording of bats is based on bat morphometrics and ultrasonic detection methods (Britzke et al., 2013). Whilst such methods provide valuable information about the bat species present, they require expensive equipment, along with practical and taxonomic expertise, and licences for the handling of bats (Hundt, 2012).

Increasingly, molecular methods are being used for identifying species, however, this requires a great deal of expertise and is not something which can be done in the field (Mayer et al., 2007, Barratt et al., 1997b). Using guano morphology to identify species requires very little by way of equipment or expertise, and can be done in the absence of bats, however, little is known about the robustness of this method.

There is little published data on identifying bat species through the use of guano morphology alone. To date, the only comprehensive identification guide for identifying Great British bats from guano was published by R.E. Stebbings in 1986 (Stebbings, 1986). This is considered the gold standard for identifying guano in the field. However, there are a number of problems faced by people using this key to identify bats in the field. Firstly, Stebbings took his measurements on fresh guano straight from the bat, however, in the field, a bat surveyor would most likely be measuring dry (old) samples, rather than samples which have been freshly deposited. Secondly, as Stebbings' guide was published in 1985, it does not cover the range of species now known to be present in Great Britain. For example, it only refers to "pipistrelle bats", and does not distinguish between *Pipistrellus pipistrellus*, *P. pygmaeus*, and *P. nathusii. P. pygmaeus* was only identified as a separate species from *P. pipistrellus* in 1997 (Barratt et al., 1997b), and *P. nathusii* was thought to be a vagrant in Great Britain until 1997 when it was found to be breeding (Barlow and Jones, 1996). Additionally, it includes data for *Myotis myotis*: since declared extinct in Great Britain. Furthermore, the data from Stebbings' guide does not indicate how deep the sampling coverage for each species was, and therefore it is not clear how reflective of the whole Great British population these measurements are.

We wished to rigorously test whether it was possible to distinguish between Great British bats species on the basis of guano morphology alone. We used mitochondrial DNA analysis order to confirm the species identity of each guano sample before taking physical measurements. This study used the parameters detailed in Stebbings' guide in addition to a

number of additional parameters, which were analysed to ascertain which of the parameters had the greatest power to distinguish between species.

6.3. Materials and methods

6.3.1. Collection of samples and mtDNA analysis

92 guano samples (each comprising multiple guanos) from 16 of the Great British bat species were collected from various locations around Great Britain. *M. alcathoe* was not included due to the low number of samples. Each sample consisted of multiple guanos (on average of 3 guanos per sample) collected at the same location and at the same time (appendix E.6.1). DNA extraction PCR, and sequencing was undertaken as described in section 3.3.1.3.

6.3.2. Selection of criterion and measurement of samples

292 dry guano samples were measured (appendix E.6.1). In order to ensure that the samples were representative of each species as possible, samples were selected to cover the whole of Great Britain (as far as the range of the species allowed). In addition to Stebbings' diagnostic characteristics of length (minimum – maximum within a sample), diameter (minimum to maximum) and particle size, we measured colour, presence/absence of nodulation, presence/absence of tip points, and presence/absence of curvature. Measurements were taken under an MX7T stereomicroscope. The criterion for categorising particle size and colour can be found in the supporting information. Nodulation, tip pointing and curvature were observed by eye and recorded as presence/absence data.

Table 6.1: Measuring colour and particle size

Particle size

1.	Very fine, smooth outline, small divots
2.	Fine, mostly smooth outline, bigger divots
3.	Medium size, rough outline, medium divots
4.	Quite coarse, rough outline, medium to large divots
5.	Coarse, very rough outline, large divots

Colour

1.	Light to dark yellow
2.	Light brown with yellow flecks present
3.	Light-medium brown
4.	Medium-dark brown with some black flecks
5.	Over half of the guano is black

6.3.3. Statistical analysis

Principle component analysis (PCA) was undertaken to compare the samples (Jolliffe, 1986). Parameters were sequentially removed from the dataset to determine the parameters with the most significant impact. Our data was also subjected to PCA with Stebbings results using just Stebbings' parameters (length, diameter and particle size). Full classical Procrustes superimposition, (orthogonal rotation, reflection, translation and scaling of one shape to find it's closest fit to another shape), and partial classical Procrustes superimposition (orthogonal rotation, reflection, translation but not scaling) were used to compare the spread of the Stebbings data set against the spread the averages of our data (Jolliffe, 1986). All Figures were created using R (R Developement Core Team, 2013).

6.4. Results

When analysing the samples using only the parameters specified in Stebbings' work (length (minimum and maximum), diameter (minimum and maximum) and particle size), there is a distinctive shift between our data set and that of Stebbings' data, figure 6.1. On average the guano that we measured was smaller than Stebbings' guano (for both length and diameter). Full Procrustes superimposition provided a better fit for our data with Stebbings' data (a Procrustes value of 11.62) than when using a partial (giving a Procrustes value of 17.93). The improvement of fit when using scaling indicates that while the overall distribution is very similar (relative positions of species similar), the distance between the mean of one species and it's morphological neighbour is greater in our dataset.

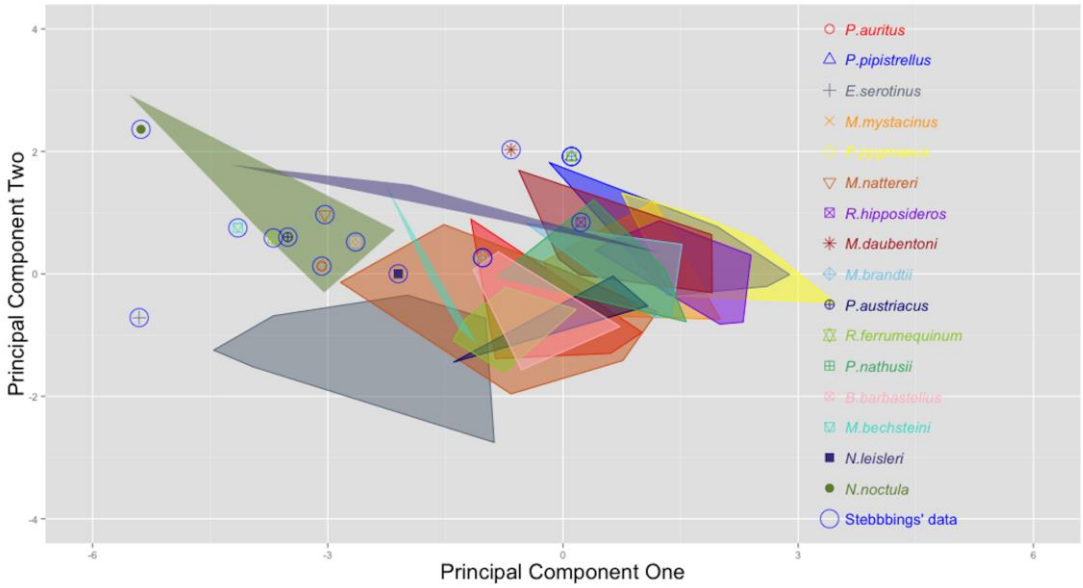


Figure 6.1: Distribution of the Great British bats in principal component space derived from the Stebbing's parameters

Principal components one and two together explain 78.6% of the variation. Points show Stebbings' data and the coloured polygons show our data. Code adapted from D.2.11.

Figure 6.2. shows that the resolution (ability to distinguish between species) is better when the curvature and tip parameters are removed, suggesting that these parameters are not useful in distinguishing bat species and merely confound the results. However, the addition of colour and nodulation (to Stebbings measurements of length, diameter and particle size) did improve the capacity to distinguish between the species.

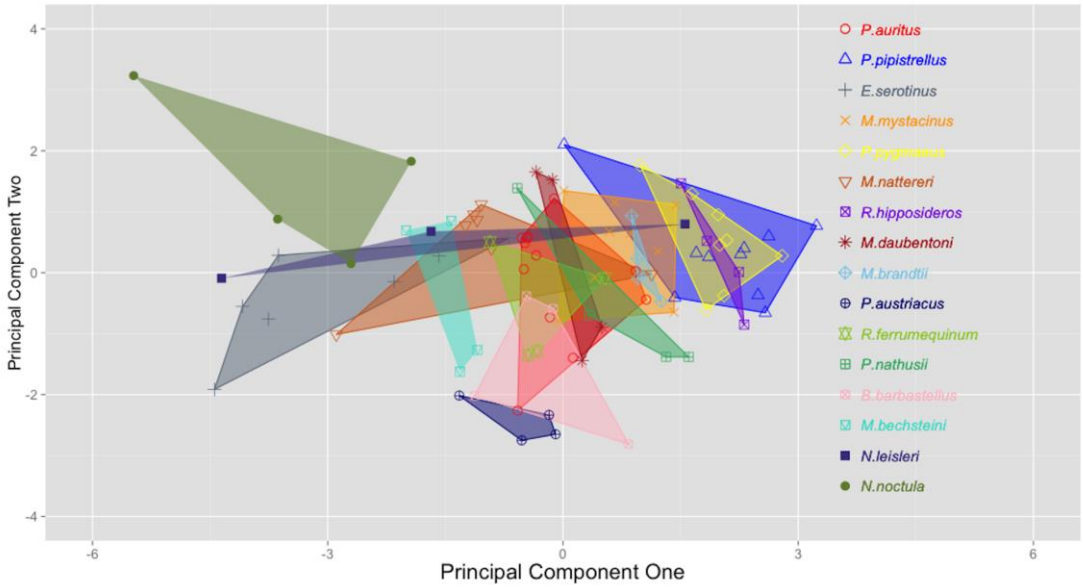


Figure 6.2: Distribution of the Great British bats in principal component space
Derived from the length, diameter, particle size, colour, and nodulation of guanos. Principal components one and two are shown; together they explain 64.7% of the variation. Where overlaps in the ranges of guano morphology can be seen, it may be possible to mistake one species for the other. The resolution (ability to distinguish between species) is finer when the "curvature" and "tip" parameters are removed, suggesting that these parameters are not useful in distinguishing bat species. Code adapted from D.2.11.

If using the final parameters (length, diameter, particle size, nodulation, and colour), some species, such as *N. noctula*, are very unlikely to be mistaken for another species (figures 6.2 + 6.3). However, species with measurements which fall within the centre of all of the Great British bat species measurements are much more likely to be wrongly identified (for example, *P. auritus* measurements overlap with 6 of the other 16 species found in Great Britain) (see table 6.1).

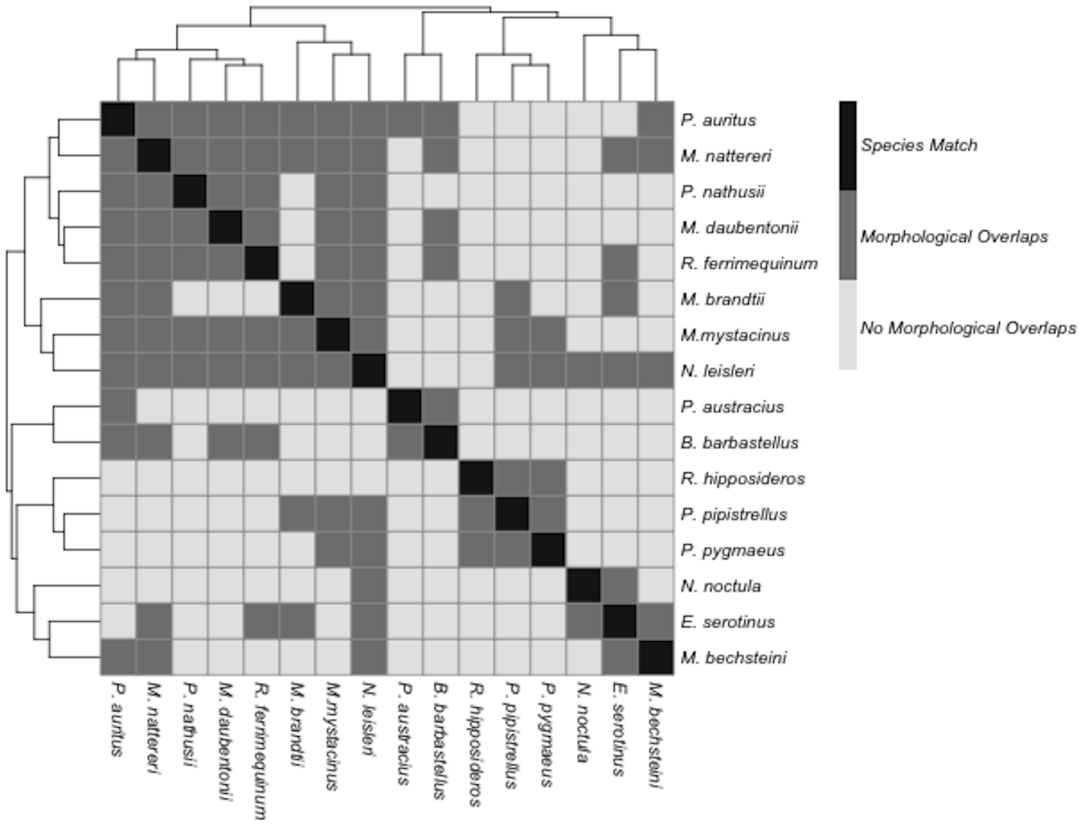


Figure 6.3: Overlaps between the guano morphology of Great British bats

Parameters measured: length, diameter, particle size, nodulation, and colour of guanos. Code adapted from D.2.4. Each species is compared pairwise to each other species to show overlaps between the morphological measurements of the guano. Where overlaps are indicated (dark grey squares), it would be possible to mistake one species for the other. Where overlaps are not present (light grey squares), it would be unlikely that one species would be mistaken for the other. Black squares are shown where a species is being compared to itself. Species are clustered by overlap similarity.

6.5. Discussion

When using guano morphology to give an indicator of bat species, it is clear that the colour of the sample, and its curvature are important parameters to measure, in addition to the traditional measurements of length, diameter and particle size (Stebbings, 1986). This is confirmed by the procrustean analyses, which showed that the means of our dataset were further apart and thus more distinct. However, what may not be captured in this data are the extremes, or where the sample comes from a juvenile, additionally, it is possible that we do not capture the range of guano colours, and thus, may be in danger of overfitting the data. As a

result, adjacent ranges, as well as overlapping ranges are likely to be areas of confusion.

There is a clear shift between our data and the data collected by Stebbings. This may be due to a number of factors. Firstly, Stebbings measured fresh guano collected directly from captive bats. It is possible that the shift is due to the guano shrinking as it dries. Alternatively, it may be due to factors such as dietary changes, perhaps due to changes in the dietary insect communities since 1986. This will be discussed in chapter 7.

6.6. Conclusions

The measurement of guano morphology may be useful where one is looking for particular species, however, its inaccuracy in most circumstances would most likely rule out this method of identifying bat species. With species normally associated with being distinct (such as *E. serotinus,* or *N. noctula*), for most of their range they are, but for areas of their range this becomes less clear. Where species have adjacent guano ranges it is unlikely that identification would be possible by eye; this would only be able to be distinguished using PCA, and would likely be of low confidence. In circumstances where the measurement of guano morphology is the only available method, it is important to ensure that the sample is dry (likely to be the case where a sample has not been freshly deposited) and that the colour and curvature of the sample are noted. The impact of diet on guano morphology will be explored in chapter 7.

Chapter 7 : Summary

7.1. Addressing the aims and hypotheses of the study

Here I will present the conclusions of the thesis about the dietary niches of each of the bat species. I will assess the extent to which each of the aims of the study have been achieved, and which of the hypotheses should be accepted.

7.1.1. Comparison of methodological approaches

As a result of the metagenome study, additional taxa were identified in the bat diets which would not likely have been amplified by the Zeale *et al* primers (section 4.3.1.1). This allowed us to develop a new system of primers which were capable of exactly matching more than twofold the number of Arthropoda sequences present on the NCBI (section 4.3.1.1). This highlights the importance of regular re-evaluation of established methods (see section 7.4.1). We have also established that using shotgun metagenome approaches can be useful for guiding primer design for metabarcoding.

The use of phylogenetic intersection analysis is of vital importance when working with next generation sequencing data. This was particularly highlighted with the Chiroptera shotgun metagenome data (section 3.4.6.2). The raw data assigned 46.97% of the Chiroptera data to the correct genus. After PIA, 90.45% of the data were assigned to the correct genus (or higher taxa consistent with the true species, as identified in section 3.3.1.3.3), see table 3.11). The success of PIA is due to its ability to assign taxa higher up the taxonomic tree based upon the support of the taxon diversity score (section 3.3.2.5.1), rather than attempting to assign all data to the terminal leaf node of the tree of life. This is demonstrated in table 5.4.

The increased Arthropoda specific data provided by metabarcoding results in an increase in the number of orders identified before and after

PIA in comparison to metagenome approaches (table 5.1). Metabarcoding returned an average of 17.6 orders before PIA, and 10.65 orders after PIA. Shotgun metagenome approaches returned an average of 7.18 orders before PIA and 7.76 orders after PIA. The increase in orders after PIA in the shotgun metagenome approach is likely due to the less complete databases for metagenome data than for metabarcode data. False assignation to model organisms occurs (often to species level). After PIA these are more robustly reassigned to higher taxonomic levels. As a result, less common taxa (or those with lower database representation) are more likely to be identified using metabarcoding approaches. Additionally, there is more data to support the presence of any given order within a guano sample using metabarcoding than using metagenome approaches.

The increased reliability of the metabarcode data in comparison to the metagenome dataset is further demonstrated through the use of Phylogenetic Intersection Analysis. Mantel tests on the full metabarcode dataset before and after PIA gives a strong observed correlation, suggesting robust support for the metabarcoding data through PIA (r= 0.99 *p*=1e-04, 9,999 permutations). Comparatively, the support for the full metagenome data by the PIA is lower (although still significant) (r=0.202, *p*= 1e-04, 9,999 permutations). The lower proportion (33.57%) of barcoding data discarded after PIA compared to the 70.4% of metagenome data discarded, further demonstrates the importance of using reference datasets with denser phylogenetic coverage. As a result, in this study, the metabarcoding data is most sensitive for identifying the range of species eaten by bats and thus measures such as niche breadth and dietary diversity are most robust using metabarcode approaches.

As reference databases are expanded, more metagenome data will be retained in analyses, increasing the efficacy of shotgun metagenome approaches in dietary analyses. This is particularly relevant as metagenome approaches have the potential to be more sensitive than metabarcoding approaches as it exploits more of the DNA laid down by

the organisms (section 3.4.8 and 3.3.2.5.2). At this stage in this project, the metagenome approach is less sensitive partly due to the patchiness of metagenome reference databases, however, if in the future full genome sequencing becomes as routine as sequencing barcoding regions, it is possible that metabarcoding approaches such as this will become redundant (see 7.4.1).

In this study, the focus has been on Arthropoda at order level. After PIA, a higher proportion of the metagenome data is assigned above order level (table 5.4) than in the metabarcode data. Indeed, in both datasets >50% of the Arthropoda data is assigned to class level (section 5.4.3). Future re-analyses of these data using more complete reference datasets would likely increase the proportions of data assigned lower down the tree for both datasets (section 7.4.1). It is also worth considering the point that as a proportion of the data returned, Arthropoda is only a small proportion in the metagenome data, compared to in the metabarcode dataset. 90.2% of the metabarcode dataset was assigned to Arthropoda, and 1.01% of the metagenome data to Arthropoda. The metabarcoding approach does return a higher overall number of orders due to the higher proportion of data returned by the metabarcode data, however, as proportion of sequencing effort (orders identified per 1,000 unique sequences), metagenome approaches return a higher number of orders (table 5.5).

With shotgun metagenome approaches, it is possible to remove artificial duplicate repeats (ADR) before analysis. This is not possible in metabarcoding data, as all metabarcoding data is, by its nature an ADR of the original template, and removal of duplicate reads in metabarcode data would remove all support for any assignment. As a result, the proportions of taxa in the metagenome data may better reflect the actual proportions in the guano. In order to minimize issues with PCR biases, each metabarcode library was comprised of six replicate PCRs.

The information provided by metagenome approaches on the other taxa which comprise the guano DNA libraries, provides a breadth of information about the bats. One example of this is the identification of non-viral pathogens within the guano (section 3.4.6.1.2). Whilst *Pseudogymnoascus destructans* (Barclay et al., 1991) (Mulec et al., 2013) was not identified in these data, metagenome methods should theoretically be capable of identifying its presence, as there is a genome sequence for *P. destructans* available (Chibucos et al., 2013).

7.1.2. Extinction risk by generalist and specialist habit

Extinction risk has been strongly linked to dietary breadth (Boyles and Storm, 2007). However, they also saw large variation within species, suggesting that diet alone cannot explain extinction risk, which is also seen here. The species identified in this study which appear to have the most specialist diets (those with narrow dietary breadths), and thus the highest predicted extinction risk were *M. alcathoe, P. nathusii, B. barbastellus, N. noctula,* and *E. serotinus* (metagenome data, after PIA), and *P. pygmaeus, R. ferrumequinum, M. nattereri, P. nathusii,* and *B. barbastellus* (metabarcode data, after PIA) (figure 5.5). The robustness of these data are discussed in 7.1.6.

The species that appear to be most robust to extinction pressures are *M. bechsteinii, N. leisleri, P. pygmaeus,* and *P. pipistrellus* (metagenome data, after PIA), and *P. auritus, M. bechsteinii, M. alcathoe,* and *N. leisleri* (metabarcode data, after PIA).

As discussed in section 7.1.1, the metabarcode data in this study is likely to be more able to identify the range of species present within the guano. As a result, the metabarcode data is likely to more useful for identifying dietary diversity and niche breadths of the bats. However, where a species is identified as vulnerable or robust by both methods of analyses (metabarcode and metagenome) we can be more confident in that identification. As, in this study, both methods analysed the same numbers of samples for both methods, where dietary diversities or niche breadths

are contradictory between methods (such as *P. pygmaeus* which is identified as vulnerable by the metabarcode data, but robust by the metagenome data), it is likely that the metabarcode data is a better reflection of the reality (*P. pygmaeus* is therefore most likely vulnerable).

The species that we predict from this study to be seemingly the most robust are *P. auritus, M. bechsteinii, M. alcathoe,* and *N. leisleri. M. bechsteinii* and *N. leisleri* were both identified as robust through both methods of analysis. *M. bechsteinii* had previously been identified as a species with a robust dietary form (figure 2.3 and 5.2) (Andreas et al., 2012a, Roswag et al., 2015, Siemers and Swift, 2006, Taake, 1993, Wolz, 1993), however, it is listed by the IUCN as declining and near threatened, suggesting that there are other factors at play.

The species that we predict to be most vulnerable are *P. pygmaeus, R. ferrumequinum, M. nattereri, P. nathusii,* and *B. barbastellus. B. barbastellus* and *P. nathusii* were identified as being vulnerable by both methods of analysis. This was supported by the data from the literature (figures 2.3 and 5.2) (Andreas et al., 2012b, Beck, 1995, Rydell et al., 1996, Sierro and Arlettaz, 1997, Whitaker Jr and Karatas, 2009, Zeale et al., 2011). *B. barbastellus* is listed by the IUCN as near threatened, with a decreasing population trend (table 1.3), which supports our findings here.

The species with narrower dietary niches appear to have earlier emergence times (table 1.1). *N. noctula* has the earliest recorded emergence time, usually emerging at around 5 minutes after sunset (Jones and Rydell, 1994), and was identified in the metagenome dataset as having a narrow dietary niche. *B. barbastellus* and *P. nathusii,* which were identified by both methods as having narrow dietary diversity, both have early emergence times compared to the other Great British species (Gelhaus and Zahn, 2010, Russo et al., 2007). Emerging earlier may be a strategy for avoiding competition with more generalist or common species. However, this does come at a price; earlier emergence time is

linked with a higher predation risk from raptorial birds, or competition by insectivorous birds (Jones and Rydell, 1994).

In the results for both datasets the species *M. bechsteinii, N. noctula, P. nathusii, R. ferrumequinum,* and *R. hipposideros,* have considerable overlaps between their diets (figure 5.6), feeding heavily on Diptera. Further work is required to assess the extent that they are feeding on the same species of Diptera, but this does appear to be the case (figure C.3.4.1). All of these species are aerial hawkers, although some have been recorded as gleaning (table 1.1). Additionally, all of these species are listed as having declining or unknown population trends (table 1.3). This intense competition between the species may be driving these population declines. Of these species, *N. noctula* and *P. nathusii* have narrow dietary diversities (as determined using both methods, figure 5.5), and thus may be the most vulnerable of these species. As both of these species have unknown population trends, these may be key species for further study.

7.1.3 Temporal and spatial dietary variation

Samples from Scotland have a high proportion of Diptera (and a low proportion of Lepidoptera) compared to samples from other regions. Species with distributions that do not extend to Scotland (*B. barbastellus, E. serotinus, M. alcathoe, M. bechsteinii, M. brandtii, P. austriacus, R. ferrumequinum,* and *R. hipposideros*), have lower average proportions of Diptera using both methods (33.68% metagenome, 47.89% metabarcode), whereas the other species have higher proportions (43.07% metagenome, 57.41% metabarcode). This is particularly marked in *M. nattereri* and *P. pipistrellus* which both have >97% of their Scottish diets based upon Diptera (compared to an average proportion of Diptera at 66.46% *M. nattereri,* and 62.26% *P. pipistrellus* throughout the rest of their ranges) (metabarcode data, appendix C.1.1). The ability to exploit Diptera as a major food source may be important in allowing species to expand their distributions to include Scotland. However, this is not the only factor: *P. auritus* have relatively low proportions of Diptera in their

diets (figure 5.4), but is well established in Scotland (appendix C.1.1), despite not having high proportions of Diptera in the Scottish diets. It is however, able to feed heavily on Diptera as seen in one of the diets from central England. As a result, it appears that in the case of *P. auritus*, opportunistic feeding, perhaps driven by variations in habitat type has driven their stable population trends (I.U.C.N., 2013).

Seasons also have a distinct impact on both the diets of the bats (section 3.5.2 and 4.5.3), and on the gut microbiota (section 3.5.3.1). This is thought to be driven by the requirements of endothermy, and by the fluctuations of available arthropods.

7.1.4. Co-habiting and cryptic species have highly diverse dietary preferences

A major mechanism of avoiding competition between cohabiting and cryptic species is the development of trophic resource partitioning (Aguirre et al., 2002, Arlettaz et al., 2000). In these data, the two generally agreed cryptic species complex (the *P. pipistrellus* and *P. pygmaeus* complex, and the small Myotis complex, section 1.6.4.1), have distinct diets (figures 5.4, 5.6). This may be as a result of similar morphological features of the cryptic species (e.g. similarities in dentition and body size), which may drive prey selection (section 7.2).

In some cases, such as that of *M. daubentonii* (figure 5.4), there are conflicting profiles between the metagenome data and the metabarcode data. This may be as a result of PCR biases causing skews in the metabarcode data (sections 4.2.1.2, 7.1.1, and 7.4.1) or due to the small samples sizes. Where highly represented orders are concerned, the metagenome data may be more robust. We therefore conclude that *M. daubentonii* is more reliant on Lepidoptera than other orders of Arthropoda.

The diets of the Yangochiroptera in comparison to the Yinpterochiroptera (figure 1.2) are not distinct from each other; both the Yinpterochiroptera have highly similar diets to some of the Yangochiroptera (figure 5.6).

7.1.5. Variations in diet do not impact guano morphology

There is a broad range in the intra-species guano measurements, and a great deal of overlap between guano morphology between species (figure 6.2). This is reflected in the breadth of dietary forms between and within species.

7.1.6. A cautionary note about sample sizes

In this study, limited numbers of samples for each species was used (6/7 per species, except *M. alcathoe* which is only represented by one guano sample). As a result, the full dietary breadth of each of the species is not captured in this study, and statistical tests lack robustness. Future work should look at each species in greater depth with increased sample numbers for each species. However, the samples and methods used have allowed us to identify some broad trends, and to compare the different methods, and assess their strengths.

7.2. Feeding styles and diets

Non-flighted (volant) arthropods in the datasets after PIA include the Amphipoda, Araneae, Archeognatha, Astigmata, Calanoida, Chordeumatida, Collembola, Cryptostigmata, Decapoda, Diplostraca, Euphausiacea, Harpacticoida, Isopoda, Ixodida, Judlia, Lithobiida, Lithobiomorpha, Mesostigmata, Mysida, Opiliones, Oribatida, Pantopoda, Pendunculata, Scorpiones, Sellilia and the Siphonaptera. Where they occur in a bat diets, they are indicators that the bat has been gleaning. Figure 7.1 shows the proportion of the diets comprised from gleaned orders.

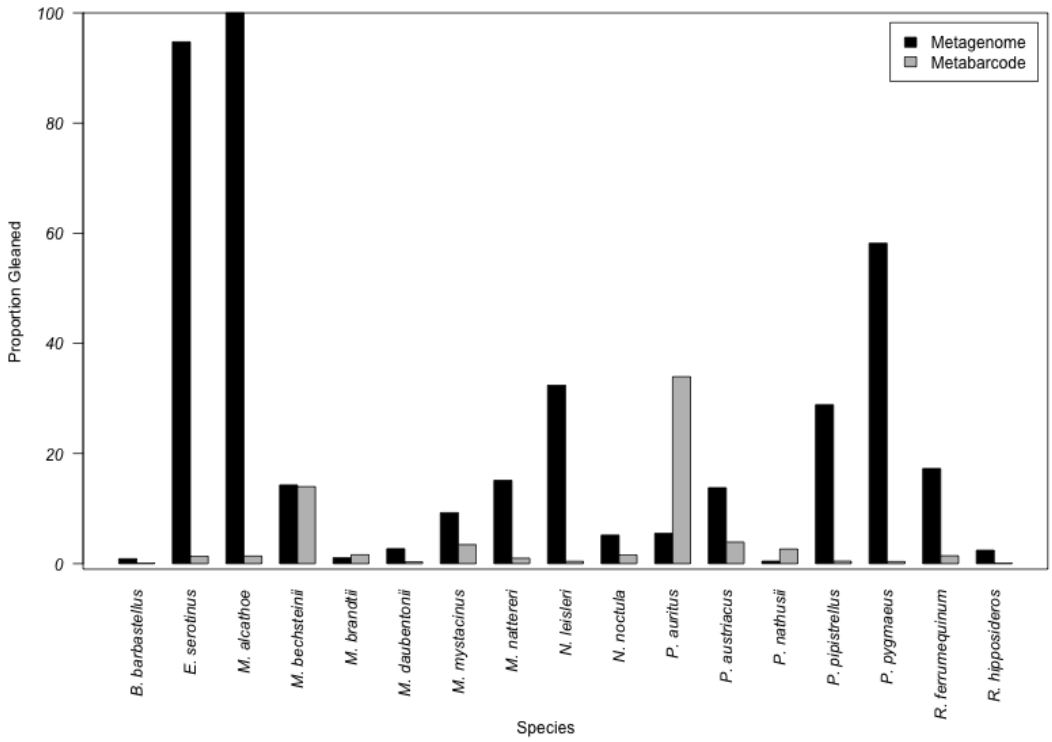


Figure 7.1: The proportions of the diet comprised by orders caught by gleaning
As described metabarcoding or shotgun metagenomics after PIA. Gleaned orders defined as any non-flighted arthropods, and includes the Amphipoda, Araneae, Archeognatha, Astigmata, Calanoida, Chordeumatida, Collembola, Cryptostigmata, Decapoda, Diplostraca, Euphausiacea, Harpacticoida, Isopoda, Ixodida, Judlia, Lithobiida, Lithobiomorpha, Mesostigmata, Mysida, Opiliones, Oribatida, Pantopoda, Pendunculata, Scorpiones, Sellilia and the Siphonaptera.

M. bechsteinii and *P. auritus* both have significant proportions of the metagenome and metabarcode datasets ascribed to gleaned orders. Both of these species have previously been described as gleaners (Coles et al., 1989, Fenton and Bogdanowicz, 2002, Petrov, 2006, Wolz, 1993).

The presence of parasitic orders in the data (including Astigmata, Cryptostigmata, and Ixodida) are thought to be indicators of grooming behavior, rather than hunting (França et al., 2013, Lourenço et al., 2014). In these data, the Astigmata are the highest represented grooming indicators in both datasets (figure 5.4). These occur in almost all bat species diet, but do not comprise a large part of the diets (except in the *M. alcathoe* metagenome data). Mesostigmata was also seen as a significant part of the *E. serotinus* diet (figure 5.4). High levels of

grooming activity may indicate poor health in the bats; *E. serotinus* showed high levels of the pathogenic bacteria *Clostridium perfringens (Hajkova and Pikula, 2007),* and *Salmonella* spp. (Arata et al., 1968, Reyes et al., 2011) (section 3.4.6.1.2).

7.3. Impacts of climate change

Global warming is predicted to increase the proportions and distributions of heat-tolerant fungal species (Garcia-Solache and Casadevall, 2010), which would narrow the thermal gradient between the ambient temperature and the T_b (body temperature), thus increasing the risk of infection (Robert and Casadevall, 2009).

Climate change also is predicted to have impact on the ranges and emergence times on insects (Cannon, 1998, Krauel et al., 2015). This will undoubtedly have impact on insectivorous mammals and birds. If ranges of arthropods extend north with warming, it is likely that the ranges of the bats may also extend northwards, particularly in the case of *B. barbastellus, E. serotinus, M. alcathoe, M. brandtii, P. austriacus, R. ferrumequinum,* and *R. hipposideros,* which are presently restricted to latitudes south of Scotland in the UK.

7.4. Niche separation of Great British bat species

We are starting to see some broad scale niche separation between species (figure 5.4). However, there is a great deal of overlap between species (figure 5.6). This is in part due to the broad range of intra-species diversities (figures 4.16 and 5.6). Some of this variation is explained by differences in season (section 4.5.3) and location (section 7.1.3). However, there is variability within these factors, which may be explained by habitat selections or by opportunistic feeding, which is particularly demonstrated in *P. auritus* (section 7.1.3).

7.5. Future directions for the field

The results of this research have posed many interesting further questions about the bat diets, outlined below. Many of these questions

can be answered through further sampling, targeted metabarcoding, further high-throughput sequencing, aDNA approaches, AMS radiocarbon dating, and climate modelling.

How similar are the diets of these species across their ranges worldwide?

By extending the sampling to other regions across the range of each bat species, further insights could be gained into the impact of environment on bat diet. In this study, in both the metabarcoding and metagenome datasets, *M. alcathoe* is under represented. In future work, additional *M. alcathoe* samples should be sourced. These were not available in this study due to the scarcity of this species (Jan et al., 2010).

How do the diets compare to the other species of insectivorous Yinpterochiroptera and Yangochiroptera?

There is great scope to extend this research across all species of bats. This would be particularly valuable in the case of species with small population sizes that are endemic to small areas such as the Florida bonneted bat, *Eumops glaucinus* (Timm and Genoways, 2004).

What is the viral load of the samples studied, and how useful is guano as a resource for non-invasive surveillance viruses in bat populations?

Undertaking library preparations to enable the sequencing of other types of viruses (not just dsDNA), could give further insight into the viral pathogens carried by bats.

How does life history stage impact diets?

Compared to other mammals, bats have unusual life histories. For their (typically) small size, bats show remarkable longevity (Jones and MacLarnon, 2001). They tend to live longer than mammals of similar size, and produce few, large offspring: a pattern seen more often in larger mammals (sometimes classified as *K* species). The age, and developmental stage, of the individual can impact its nutritional

requirements and diets often vary throughout the life cycle of the organism (Troyer, 1984). This is particularly the case with juveniles or during pregnancy and lactation (Kunz et al., 1995, Rydell, 1989a). This could be studied using captured bats if life history stage and sex is recorded. These can be held until guano has been deposited, then released. Whilst this may not give exact ages of the bats, activities such as ringing may help to resolve this. Long term captive bats (those not able to be released into the wild) could be used as a source of control diets.

How does bat morphology impact diet?

We are starting to see intra-species diet variation emerging, most notably in the cases of cryptic species such as the *P. pipistrellus/P. pygmaeus* complex and the small myotis complex (section 7.1.4). It would be useful to examine these data in the context of traits such as dentition, jaw size, and other morphological features. Durophagy and bite force, especially in Myotis, is thought to be a driver in diet selection (Ghazali and Dzeverin, 2013, Gill et al., 2014, Freeman, 1998, Freeman, 1979, Freeman and Lemen, 2007, Krüger et al., 2014). This was not done here due to the lack of metadata on life history.

How does heterothermy impact bacterial and fungal flora?

Whilst the metagenome data has gone some way to describing the impacts of heterothermy on the gut microbiota, further information could be gained using a targeted metabarcoding approach. Additionally, further sampling of single sites across a time series would help to remove confounding factors.

Do bat diets reflect long-term environmental changes?

Guano deposits in caves have been used to study environmental changes (Bird et al., 2007, Shahack-Gross et al., 2004). There is precedent for studying aDNA from cave sites (Briggs et al., 2007), along with recent developments in the processing and analysis of metagenomic DNA from archaeological samples (Kistler et al., 2015, Smith et al.,

2015). We suggest there is scope for using cave guano deposits to study dietary change over time using molecular methods.

What would be the impact of changing arthropod distributions driven by habitat development and climate change on bat diets?
Climate change is predicted to have impact on the ranges and emergence times on insects (Cannon, 1998, Krauel et al., 2015). Both of these factors may have significant consequences on bats which rely on these. Changes in prey density have been shown to have direct impact on bat populations and distributions (Charbonnier et al., 2014). Projections of the impact of climate change on arthropod availability should be undertaken in order to identify areas of vulnerability in bat populations.

7.5.1. Bioinformatic considerations for future work

There are a number of technical considerations which we advise investigating before future work. The advancements in sequencing technologies has been fundamental in allowing the development of this research, however it is not without issues. PCR bias is a key drawback, particularly in the case of the metabarcode data, which has been subjected to more rounds of PCR, and for which artificial duplicate reads are not removed. For the metagenome data, single molecule sequencing on platforms such as Oxford Nanopore (Mikheyev and Tin, 2014) or PacBio (Quail et al., 2012) would help to avoid this, as would Illumina PCR free library preparation methods (Kozarewa et al., 2009). Single molecule sequencing could be useful where using degraded samples, however, due to the low endogenous DNA content of bat guano and the relatively low output of currently available single molecule sequencers, these may not be useful for sequencing the Arthropoda fraction of the DNA. PCR free methods for Illumina may be more useful due to their high data yields.

It is unlikely that the primer set designed will be able to capture the full breadth of the Arthropoda, as the majority of Arthropoda CO1 genes have

not been sequenced. As of April 2016 103,503 species were represented on the BOLD database (Ratnasingham and Hebert, 2007), whereas there are ~6.8 million species (range 5.9-7.8 million) predicted terrestrial Arthropoda species globally (Stork et al., 2015). As sequence databases are continually updated, it is advisable that primers used for metabarcoding are regularly assessed and, if necessary, re-designed. In this dataset, there were only 72 sequences matched primer >18_F_Arthropoda_COI, suggesting that it is not useful, and should not be used in future. Additionally, primer >11_R_Arthropoda_COI amplified high proportions (90.16%) of spurious sequences, and should also be excluded from future work.

A large proportion (42.11%) of the metagenome data was discarded due to lack of matches on the reference databases. As database coverage increases in future, re-analyses of metagenome data may improve sensitivity towards rarer species (Ostell et al., 2004).

Appendices

Appendix A. Lab Protocols

A.1. Technical considerations when working with low quality DNA

When working with ancient DNA (aDNA) or DNA with low copy number, contamination is a key issue, particularly from amplified DNA from PCR. It is therefore important that dedicated lab spaces are used, as follows:

Ancient DNA lab: This is a chambered lab used solely for the extraction of DNA from archaeological/historical material and sample preparation prior to PCR.

Modern DNA lab: This lab is used for DNA extraction of modern material and sample preparation prior to PCR.

Post-PCR lab: This lab is used for PCR and all post-PCR activities and has communal fume hoods and gel electrophoresis/ imaging facilities, which we share with other groups.

It is important to avoid moving equipment and reagents between labs. This is strictly forbidden in the clean lab and best avoided in the modern labs. Lab coats are kept within their own separate labs. PCR products are never taken into extraction labs. All equipment and surfaces are wiped with bleach followed by 70% ethanol after use. Racks are soaked in a bleach bath after use, then rinsed with 70% ethanol. Gloves and booties are disposed immediately after taking them off.

The aDNA lab must never be entered after being in any other lab. Before entering the aDNA lab, freshly laundered clothes must be worn, and all lab users must have showered first. No personal items (e.g. lab books, laptops) must be taken into the aDNA lab. The following must always be worn: coveralls (e.g. Microguard), overshoes, face mask, hair net, two pairs of gloves (the inside pair not to be removed inside the lab).

A.2. Homebrew SPRI beads

Homebrew SPRI beads are used as they are around ~98% cheaper than Ampure beads. Based on the recipe from (Rohland and Reich, 2012). The addition of tween helps movement of beads to the magnet.

- Add 9g PEG-8000, 10 mL 5 M NaCl, 500 uL 1 M Tris-HCl, and 100 uL 0.5 M EDTA to a 50ml tube.
- Top up to 49 mL with water, and shake until all PEG is dissolved.
- Add 27.5 uL Tween-20 and shake.
- Shake Sera-Mag beads to re-suspend completely, and transfer 1mL bead suspension to a 1.5 tube. Pellet beads on a magnetic stand and discard supernatant. Wash twice with 1 mL TE buffer each time, and then re-suspend in 1 mL TE buffer.
- Add beads to the PEG solution and mix.
- Store at 4°C in the dark.

A.3. EBT buffer recipe

This is identical to QIAGEN's EB buffer, with the addition of tween, which helps movement of beads in SPRI bead clean-up. This is used throughout the library preparation protocols in chapters three and four.

10 mM Tris-Cl, pH should be 8.0 and 8.5
0.05% Tween 20

A.4. Oligonucleotide hybridisation buffer

This is used to hybridise the adaptor oligonucleotides used in the library preparation protocols in chapters three and four. Taken from (Meyer and Kircher, 2010)

500 mM NaCl
10 mM Tris-Cl, pH 8.0
1 mM EDTA, pH 8.0

Appendix B: Sample sheets

All sample sheets have been summarised to show only completed cells

Chapter 3:

B.3.1. MiSeq run 1

[Header]				
IEMFileVersion	4			
Investigator Name	Roselyn Ware			
Experiment Name	Rosie_miseq2			
Date	08/06/2014			
Workflow	GenerateFASTQ			
Application	FASTQ Only			
Assay	Nextera XT			
Description	Miseq1			
Chemistry	Amplicon			
[Reads]				
251				
251				
[Settings]				
ReverseComplement	0			
Adapter	CTGTCTCTTATACACATCT			
[Data]				
Sample_ID	I7_Index_ID	index	I5_Index_ID	index2
1. *P. auritus*	N701	CGAGTAAT	S501	TATAGCCT
2. *P. auritus*	N702	CGAGTAAT	S501	ATAGAGGC
3. *P. auritus*	N703	CGAGTAAT	S501	CCTATCCT
4. *P. auritus*	N704	CGAGTAAT	S501	GGCTCTGA
5. *P. auritus*	N705	CGAGTAAT	S501	AGGCGAAG
6. *P. pipistrellus*	N706	CGAGTAAT	S501	TAATCTTA
7. *P. pipistrellus*	N707	CGAGTAAT	S501	CAGGACGT
8. *P. pipistrellus*	N708	CGAGTAAT	S501	TGACTGAC
9. *P. pipistrellus*	N701	TCTCCGGA	S502	TATAGCCT
10. *P. pipistrellus*	N702	TCTCCGGA	S502	ATAGAGGC
11. *P. pipistrellus*	N703	TCTCCGGA	S502	CCTATCCT
12. *E. serotinus*	N704	TCTCCGGA	S502	GGCTCTGA
13. *E. serotinus*	N705	TCTCCGGA	S502	AGGCGAAG
14. *E. serotinus*	N706	TCTCCGGA	S502	TAATCTTA
15. *E. serotinus*	N707	TCTCCGGA	S502	CAGGACGT

16. *M. mystacinus*	N708	TCTCCGGA	S502	TGACTGAC
17. *M. mystacinus*	N701	AATGAGCG	S503	TATAGCCT
18. *M. mystacinus*	N702	AATGAGCG	S503	ATAGAGGC
19. *M. mystacinus*	N703	AATGAGCG	S503	CCTATCCT
20. *P. pygmaeus*	N704	AATGAGCG	S503	GGCTCTGA
21. *P. pygmaeus*	N705	AATGAGCG	S503	AGGCGAAG
22. *P. pygmaeus*	N706	AATGAGCG	S503	TAATCTTA
23. *P. pygmaeus*	N707	AATGAGCG	S503	CAGGACGT
24. *M. nattereri*	N708	AATGAGCG	S503	TGACTGAC
25. *M. nattereri*	N701	GGAATCTC	S504	TATAGCCT
26. *M. nattereri*	N702	GGAATCTC	S504	ATAGAGGC
27. *M. nattereri*	N703	GGAATCTC	S504	CCTATCCT
28. *M. nattereri*	N704	GGAATCTC	S504	GGCTCTGA
29. *R. hipposideros*	N705	GGAATCTC	S504	AGGCGAAG
30. *R. hipposideros*	N706	GGAATCTC	S504	TAATCTTA
31. *R. hipposideros*	N707	GGAATCTC	S504	CAGGACGT
32. *R. hipposideros*	N708	GGAATCTC	S504	TGACTGAC
33. *M. daubentonii*	N701	TTCTGAAT	S505	TATAGCCT
34. *M. daubentonii*	N702	TTCTGAAT	S505	ATAGAGGC
35. *M. daubentonii*	N703	TTCTGAAT	S505	CCTATCCT
36. *M. daubentonii*	N704	TTCTGAAT	S505	GGCTCTGA
37. *M. brandtii*	N705	TTCTGAAT	S505	AGGCGAAG
38. *M. brandtii*	N706	TTCTGAAT	S505	TAATCTTA
39. *M. brandtii*	N707	TTCTGAAT	S505	CAGGACGT
40. *M. brandtii*	N708	TTCTGAAT	S505	TGACTGAC
41. *P. austriacus*	N701	ACGAATTC	S506	TATAGCCT
42. *P. austriacus*	N702	ACGAATTC	S506	ATAGAGGC
43. *P. austriacus*	N703	ACGAATTC	S506	CCTATCCT
44. *R. ferrumequinum*	N704	ACGAATTC	S506	GGCTCTGA
45. *R. ferrumequinum*	N705	ACGAATTC	S506	AGGCGAAG
46. *R. ferrumequinum*	N706	ACGAATTC	S506	TAATCTTA
47. *R. ferrumequinum*	N707	ACGAATTC	S506	CAGGACGT
48. *P. nathusii*	N708	ACGAATTC	S506	TGACTGAC
49. *P. nathusii*	N701	AGCTTCAG	S507	TATAGCCT
50. *P. nathusii*	N702	AGCTTCAG	S507	ATAGAGGC
51. *P. nathusii*	N703	AGCTTCAG	S507	CCTATCCT
52. *B. barbastellus*	N704	AGCTTCAG	S507	GGCTCTGA
53. *B. barbastellus*	N705	AGCTTCAG	S507	AGGCGAAG
54. *B. barbastellus*	N706	AGCTTCAG	S507	TAATCTTA

55. *B. barbastellus*	N707	AGCTTCAG	S507	CAGGACGT
56. *M. bechsteinii*	N708	AGCTTCAG	S507	TGACTGAC
57. *M. bechsteinii*	N701	GCGCATTA	S508	TATAGCCT
58. *M. bechsteinii*	N702	GCGCATTA	S508	ATAGAGGC
59. *N. leisleri*	N703	GCGCATTA	S508	CCTATCCT
60. *N. leisleri*	N704	GCGCATTA	S508	GGCTCTGA
61. *N. noctula*	N705	GCGCATTA	S508	AGGCGAAG
62. *N. noctula*	N706	GCGCATTA	S508	TAATCTTA
63. Positive control	N707	GCGCATTA	S508	CAGGACGT
64. *N. leisleri*	N708	GCGCATTA	S508	TGACTGAC
65. *N. leisleri*	N701	CATAGCCG	S509	TATAGCCT
66. *N. noctula*	N702	CATAGCCG	S509	ATAGAGGC
67. *M. bechsteinii*	N703	CATAGCCG	S509	CCTATCCT
68. Blank	N704	CATAGCCG	S509	GGCTCTGA
69. *P. auritus*	N705	CATAGCCG	S509	AGGCGAAG
70. *P. austriacus*	N706	CATAGCCG	S509	TAATCTTA
71. *P. auritus*	N707	CATAGCCG	S509	CAGGACGT

B.3.2. MiSeq run 2

[Header]				
IEMFileVersion	4			
Investigator Name	Roselyn ware			
Experiment Name	Rosie_Meta2			
Date	07/06/2015			
Workflow	GenerateFASTQ			
Application	FASTQ Only			
Assay	Nextera XT			
Description	Miseq			
Chemistry	Amplicon			
[Reads]				
251				
251				
[Settings]				
ReverseComplement	0			
Adapter	CTGTCTCTTATACACATCT			
[Data]				
Sample_ID	I7_Index_ID	index	I5_Index_ID	index2
1. *B. barbastellus*	M701	CGAGTAAT	T501	TATAGCCT
2. *P. pipistrellus*	M702	TCTCCGGA	T501	TATAGCCT

3. *B. barbastellus*	M703	AATGAGCG	T501	TATAGCCT
4. *M. nattereri*	M704	GGAATCTC	T501	TATAGCCT
5. *B. barbastellus*	M705	TTCTGAAT	T501	TATAGCCT
6. *B. barbastellus*	M706	ACGAATTC	T501	TATAGCCT
7. *E. serotinus*	M707	AGCTTCAG	T501	TATAGCCT
8. *M. bechsteinii*	M708	GCGCATTA	T501	TATAGCCT
9. *P. auritus*	M701	CGAGTAAT	T502	ATAGAGGC
10. *P. pipistrellus*	M702	TCTCCGGA	T502	ATAGAGGC
11. *E. serotinus*	M703	AATGAGCG	T502	ATAGAGGC
12. *M. nattereri*	M704	GGAATCTC	T502	ATAGAGGC
13. *M. daubentonii*	M705	TTCTGAAT	T502	ATAGAGGC
14. *P. austriacus*	M706	ACGAATTC	T502	ATAGAGGC
15. *E. serotinus*	M707	AGCTTCAG	T502	ATAGAGGC
16. *E. serotinus*	M708	GCGCATTA	T502	ATAGAGGC
17. *E. serotinus*	M701	CGAGTAAT	T503	CCTATCCT
18. *E. serotinus*	M702	TCTCCGGA	T503	CCTATCCT
19. *M. alcathoe*	M703	AATGAGCG	T503	CCTATCCT
20. *M. bechsteinii*	M704	GGAATCTC	T503	CCTATCCT
21. *M. bechsteinii*	M705	TTCTGAAT	T503	CCTATCCT
22. *M. bechsteinii*	M706	ACGAATTC	T503	CCTATCCT
23. *M. bechsteinii*	M707	AGCTTCAG	T503	CCTATCCT
24. *N. leisleri*	M708	GCGCATTA	T503	CCTATCCT
25. *M. brandtii*	M701	CGAGTAAT	T504	GGCTCTGA
26. *M. brandtii*	M702	TCTCCGGA	T504	GGCTCTGA
27. *M. brandtii*	M703	AATGAGCG	T504	GGCTCTGA
28. *M. brandtii*	M704	GGAATCTC	T504	GGCTCTGA
29. *M. daubentonii*	M705	TTCTGAAT	T504	GGCTCTGA
30. *M. daubentonii*	M706	ACGAATTC	T504	GGCTCTGA
31. *M. daubentonii*	M707	AGCTTCAG	T504	GGCTCTGA
32. *M. daubentonii*	M708	GCGCATTA	T504	GGCTCTGA
33. *M. daubentonii*	M701	CGAGTAAT	T505	AGGCGAAG
34. *M. daubentonii*	M702	TCTCCGGA	T505	AGGCGAAG
35. *M. mystacinus*	M703	AATGAGCG	T505	AGGCGAAG
36. *M. mystacinus*	M704	GGAATCTC	T505	AGGCGAAG
37. *M. brandtii*	M705	TTCTGAAT	T505	AGGCGAAG
38. *M. mystacinus*	M706	ACGAATTC	T505	AGGCGAAG
39. *M. mystacinus*	M707	AGCTTCAG	T505	AGGCGAAG
40. *M. mystacinus*	M708	GCGCATTA	T505	AGGCGAAG
41. *M. mystacinus*	M701	CGAGTAAT	T506	TAATCTTA

42. *M. mystacinus*	M702	TCTCCGGA	T506	TAATCTTA
43. *M. nattereri*	M703	AATGAGCG	T506	TAATCTTA
44. *M. nattereri*	M704	GGAATCTC	T506	TAATCTTA
45. *M. brandtii*	M705	TTCTGAAT	T506	TAATCTTA
46. *M. nattereri*	M706	ACGAATTC	T506	TAATCTTA
47. *B. barbastellus*	M707	AGCTTCAG	T506	TAATCTTA
48. *M. nattereri*	M708	GCGCATTA	T506	TAATCTTA
49. *P. pipistrellus*	M701	CGAGTAAT	T507	CAGGACGT
50. *M. nattereri*	M702	TCTCCGGA	T507	CAGGACGT
51. *N. leisleri*	M703	AATGAGCG	T507	CAGGACGT
52. *N. leisleri*	M704	GGAATCTC	T507	CAGGACGT
53. *N. leisleri*	M705	TTCTGAAT	T507	CAGGACGT
54. *N. leisleri*	M706	ACGAATTC	T507	CAGGACGT
55. *B. barbastellus*	M707	AGCTTCAG	T507	CAGGACGT
56. *N. leisleri*	M708	GCGCATTA	T507	CAGGACGT
57. *N. noctula*	M701	CGAGTAAT	T508	TGACTGAC
58. *N. noctula*	M702	TCTCCGGA	T508	TGACTGAC
59. *N. noctula*	M703	AATGAGCG	T508	TGACTGAC
60. *N. noctula*	M704	GGAATCTC	T508	TGACTGAC
61. *M. brandtii*	M705	TTCTGAAT	T508	TGACTGAC
62. *N. noctula*	M706	ACGAATTC	T508	TGACTGAC
63. *N. noctula*	M707	AGCTTCAG	T508	TGACTGAC
64. *P. auritus*	M708	GCGCATTA	T508	TGACTGAC
65. *P. auritus*	M701	CGAGTAAT	T509	GTCACATG
66. *P. auritus*	M702	TCTCCGGA	T509	GTCACATG
67. *P. auritus*	M703	AATGAGCG	T509	GTCACATG
68. *P. auritus*	M704	GGAATCTC	T509	GTCACATG
69. *P. austriacus*	M705	TTCTGAAT	T509	GTCACATG
70. *P. austriacus*	M706	ACGAATTC	T509	GTCACATG
71. *P. austriacus*	M707	AGCTTCAG	T509	GTCACATG
72. *P. austriacus*	M708	GCGCATTA	T509	GTCACATG
73. *P. austriacus*	M701	CGAGTAAT	T510	ACTGTACG
74. *P. austriacus*	M702	TCTCCGGA	T510	ACTGTACG
75. *P. nathusii*	M703	AATGAGCG	T510	ACTGTACG
76. *P. nathusii*	M704	GGAATCTC	T510	ACTGTACG
77. *P. nathusii*	M705	TTCTGAAT	T510	ACTGTACG
78. *P. nathusii*	M706	ACGAATTC	T510	ACTGTACG
79. *P. nathusii*	M707	AGCTTCAG	T510	ACTGTACG
80. *P. nathusii*	M708	GCGCATTA	T510	ACTGTACG

81. *P. pipistrellus*	M701	CGAGTAAT	T511	GCGCATTC
82. *P. pipistrellus*	M702	TCTCCGGA	T511	GCGCATTC
83. *P. pipistrellus*	M703	AATGAGCG	T511	GCGCATTC
84. *P. pipistrellus*	M704	GGAATCTC	T511	GCGCATTC
85. *P. pygmaeus*	M705	TTCTGAAT	T511	GCGCATTC
86. *P. pygmaeus*	M706	ACGAATTC	T511	GCGCATTC
87. *P. pygmaeus*	M707	AGCTTCAG	T511	GCGCATTC
88. *P. pygmaeus*	M708	GCGCATTA	T511	GCGCATTC
89. *P. pygmaeus*	M701	CGAGTAAT	T512	CTCTGGAA
90. Positive control	M702	TCTCCGGA	T512	CTCTGGAA
91. *P. pygmaeus*	M703	AATGAGCG	T512	CTCTGGAA
92. *P. pygmaeus*	M704	GGAATCTC	T512	CTCTGGAA
93. *R. ferrumequinum*	M705	TTCTGAAT	T512	CTCTGGAA
94. *R. ferrumequinum*	M706	ACGAATTC	T512	CTCTGGAA
95. *R. ferrumequinum*	M707	AGCTTCAG	T512	CTCTGGAA
96. *R. ferrumequinum*	M708	GCGCATTA	T512	CTCTGGAA
97. *R. ferrumequinum*	M709	CATAGCCG	T501	TATAGCCT
98. *R. ferrumequinum*	M709	CATAGCCG	T502	ATAGAGGC
99. *M. bechsteinii*	M709	CATAGCCG	T503	CCTATCCT
100. *R. hipposideros*	M709	CATAGCCG	T504	GGCTCTGA
101. *R. hipposideros*	M709	CATAGCCG	T505	AGGCGAAG
102. *P. auritus*	M709	CATAGCCG	T506	TAATCTTA
103. *R. hipposideros*	M709	CATAGCCG	T508	CAGGACGT
104. *R. hipposideros*	M709	CATAGCCG	T507	TGACTGAC
105. *R. hipposideros*	M709	CATAGCCG	T509	GTCACATG
106. *R. hipposideros*	M709	CATAGCCG	T510	ACTGTACG
107. *R. hipposideros*	M709	CATAGCCG	T511	GCGCATTC
108. Blank	M709	CATAGCCG	T512	CTCTGGAA

Chapter 4:

B.4.1. MiSeq run

[Header]				
IEMFileVersion	4			
Investigator Name	Roselyn Ware			
Experiment Name	Rosie_miseq_ArtCOI			
Date	24/02/2015			
Workflow	GenerateFASTQ			
Application	FASTQ Only			
Assay	Nextera XT			

Description	Miseq1			
Chemistry	Amplicon			
[Reads]				
251				
251				
[Settings]				
ReverseComplement	0			
Adapter	CTGTCTCTTATACACATCT			
[Data]				
Sample_ID	I7_Index_ID	index	I5_Index_ID	index2
1. *B. barbastellus*	N701	CGAGTAAT	S501	TATAGCCT
2. *M. bechsteinii*	N702	TCTCCGGA	S501	TATAGCCT
3. *M. daubentonii*	N703	AATGAGCG	S501	TATAGCCT
4. *M. nattereri*	N704	GGAATCTC	S501	TATAGCCT
5. *N. noctula*	N705	TTCTGAAT	S501	TATAGCCT
6. *P. austriacus*	N706	ACGAATTC	S501	TATAGCCT
7. *P. pipistrellus*	N707	AGCTTCAG	S501	TATAGCCT
8. *R. ferrumequinum*	N708	GCGCATTA	S501	TATAGCCT
9. *B. barbastellus*	N701	CGAGTAAT	S502	ATAGAGGC
10. *M. bechsteinii*	N702	TCTCCGGA	S502	ATAGAGGC
11. *M. daubentonii*	N703	AATGAGCG	S502	ATAGAGGC
12. *M. nattereri*	N704	GGAATCTC	S502	ATAGAGGC
13. *N. noctula*	N705	TTCTGAAT	S502	ATAGAGGC
14. *P. austriacus*	N706	ACGAATTC	S502	ATAGAGGC
15. *P. pipistrellus*	N707	AGCTTCAG	S502	ATAGAGGC
16. *R. ferrumequinum*	N708	GCGCATTA	S502	ATAGAGGC
17. *B. barbastellus*	N701	CGAGTAAT	S503	CCTATCCT
18. *M. bechsteinii*	N702	TCTCCGGA	S503	CCTATCCT
19. *M. daubentonii*	N703	AATGAGCG	S503	CCTATCCT
20. *M. nattereri*	N704	GGAATCTC	S503	CCTATCCT
21. *N. noctula*	N705	TTCTGAAT	S503	CCTATCCT
22. *P. austriacus*	N706	ACGAATTC	S503	CCTATCCT
23. *P. pipistrellus*	N707	AGCTTCAG	S503	CCTATCCT
24. *R. ferrumequinum*	N708	GCGCATTA	S503	CCTATCCT
25. *B. barbastellus*	N701	CGAGTAAT	S504	GGCTCTGA
26. *M. bechsteinii*	N702	TCTCCGGA	S504	GGCTCTGA
27. *M. daubentonii*	N703	AATGAGCG	S504	GGCTCTGA
28. *M. nattereri*	N704	GGAATCTC	S504	GGCTCTGA
29. *N. noctula*	N705	TTCTGAAT	S504	GGCTCTGA

30. *P. austriacus*	N706	ACGAATTC	S504	GGCTCTGA
31. *P. pipistrellus*	N707	AGCTTCAG	S504	GGCTCTGA
32. *R. ferrumequinum*	N708	GCGCATTA	S504	GGCTCTGA
33. *B. barbastellus*	N701	CGAGTAAT	S505	AGGCGAAG
34. *M. bechsteinii*	N702	TCTCCGGA	S505	AGGCGAAG
35. *M. daubentonii*	N703	AATGAGCG	S505	AGGCGAAG
36. *M. nattereri*	N704	GGAATCTC	S505	AGGCGAAG
37. *N. noctula*	N705	TTCTGAAT	S505	AGGCGAAG
38. *P. austriacus*	N706	ACGAATTC	S505	AGGCGAAG
39. *P. pipistrellus*	N707	AGCTTCAG	S505	AGGCGAAG
40. *R. ferrumequinum*	N708	GCGCATTA	S505	AGGCGAAG
41. *B. barbastellus*	N701	CGAGTAAT	S506	TAATCTTA
42. *M. brandtii*	N702	TCTCCGGA	S506	TAATCTTA
43. *M. mystacinus*	N703	AATGAGCG	S506	TAATCTTA
44. *N. leisleri*	N704	GGAATCTC	S506	TAATCTTA
45. *P. auritus*	N705	TTCTGAAT	S506	TAATCTTA
46. *P. nathusii*	N706	ACGAATTC	S506	TAATCTTA
47. *P. pygmaeus*	N707	AGCTTCAG	S506	TAATCTTA
48. *R. hipposideros*	N708	GCGCATTA	S506	TAATCTTA
57. *E. serotinus*	N701	CGAGTAAT	S507	CAGGACGT
58. *M. brandtii*	N702	TCTCCGGA	S507	CAGGACGT
59. *M. mystacinus*	N703	AATGAGCG	S507	CAGGACGT
60. *N. leisleri*	N704	GGAATCTC	S507	CAGGACGT
61. *P. auritus*	N705	TTCTGAAT	S507	CAGGACGT
62. *P. nathusii*	N706	ACGAATTC	S507	CAGGACGT
63. *P. pygmaeus*	N707	AGCTTCAG	S507	CAGGACGT
64. *R. hipposideros*	N708	GCGCATTA	S507	CAGGACGT
49. *E. serotinus*	N701	CGAGTAAT	S508	TGACTGAC
50. *M. brandtii*	N702	TCTCCGGA	S508	TGACTGAC
51. *M. mystacinus*	N703	AATGAGCG	S508	TGACTGAC
52. *N. leisleri*	N704	GGAATCTC	S508	TGACTGAC
53. *P. auritus*	N705	TTCTGAAT	S508	TGACTGAC
54. *P. nathusii*	N706	ACGAATTC	S508	TGACTGAC
55. *P. pygmaeus*	N707	AGCTTCAG	S508	TGACTGAC
56. *R. hipposideros*	N708	GCGCATTA	S508	TGACTGAC
65. *E. serotinus*	N701	CGAGTAAT	S509	GTCACATG
66. *M. brandtii*	N702	TCTCCGGA	S509	GTCACATG
67. *M. mystacinus*	N703	AATGAGCG	S509	GTCACATG
68. *N. leisleri*	N704	GGAATCTC	S509	GTCACATG

69. *P. auritus*	N705	TTCTGAAT	S509	GTCACATG
70. *P. nathusii*	N706	ACGAATTC	S509	GTCACATG
71. *P. pygmaeus*	N707	AGCTTCAG	S509	GTCACATG
72. *R. hipposideros*	N708	GCGCATTA	S509	GTCACATG
73. *E. serotinus*	N701	CGAGTAAT	S510	ACTGTACG
74. *M. brandtii*	N702	TCTCCGGA	S510	ACTGTACG
75. *M. mystacinus*	N703	AATGAGCG	S510	ACTGTACG
76. *N. leisleri*	N704	GGAATCTC	S510	ACTGTACG
77. *P. auritus*	N705	TTCTGAAT	S510	ACTGTACG
78. *P. nathusii*	N706	ACGAATTC	S510	ACTGTACG
79. *P. pygmaeus*	N707	AGCTTCAG	S510	ACTGTACG
80. *R. hipposideros*	N708	GCGCATTA	S510	ACTGTACG
81. *E. serotinus*	N701	CGAGTAAT	S511	GCGCATTC
82. *M. brandtii*	N702	TCTCCGGA	S511	GCGCATTC
83. *M. mystacinus*	N703	AATGAGCG	S511	GCGCATTC
84. *N. leisleri*	N704	GGAATCTC	S511	GCGCATTC
85. *P. auritus*	N705	TTCTGAAT	S511	GCGCATTC
86. *P. nathusii*	N706	ACGAATTC	S511	GCGCATTC
87. *P. pygmaeus*	N707	AGCTTCAG	S511	GCGCATTC
88. *R. hipposideros*	N708	GCGCATTA	S511	GCGCATTC
89. *E. serotinus*	N701	CGAGTAAT	S512	CTCTGGAA
90. *M. daubentonii*	N702	TCTCCGGA	S512	CTCTGGAA
91. *M. nattereri*	N703	AATGAGCG	S512	CTCTGGAA
92. *N. noctula*	N704	GGAATCTC	S512	CTCTGGAA
93. *P. austriacus*	N705	TTCTGAAT	S512	CTCTGGAA
94. *P. pipistrellus*	N706	ACGAATTC	S512	CTCTGGAA
95. *R. ferrumequinum*	N707	AGCTTCAG	S512	CTCTGGAA
96. Blank	N708	GCGCATTA	S512	CTCTGGAA
97. *M. alcathoe*	N709	CATAGCCG	S501	TATAGCCT
98. *M. bechsteinii*	N709	CATAGCCG	S502	ATAGAGGC
99. *M. mystacinus*	N709	CATAGCCG	S503	CCTATCCT
100. *P. auritus*	N709	CATAGCCG	S504	GGCTCTGA
101. *P. austriacus*	N709	CATAGCCG	S505	AGGCGAAG
102. *P. pipistrellus*	N709	CATAGCCG	S506	TAATCTTA
103. *P. pygmaeus*	N709	CATAGCCG	S507	TGACTGAC
104. *P. pygmaeus*	N709	CATAGCCG	S508	CAGGACGT
105. *M. daubentonii*	N709	CATAGCCG	S509	GTCACATG
106. *M. nattereri*	N709	CATAGCCG	S510	ACTGTACG
107. *R. hipposideros*	N709	CATAGCCG	S511	GCGCATTC

108. *M. brandtii*	N709	CATAGCCG	S512	CTCTGGAA

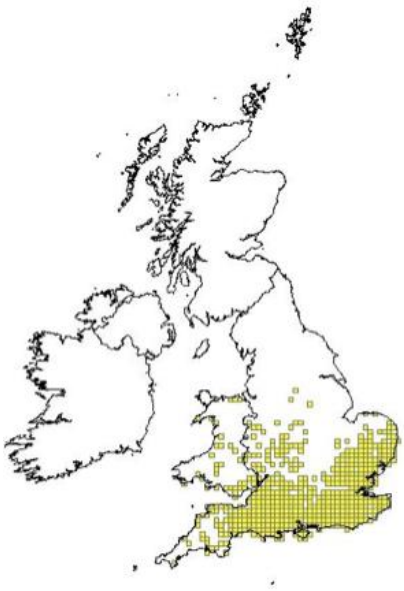
Appendix C: Figures

Chapter 1:

C.1.1. The distributions of great British bat species. Adapted from the National Bat Monitoring Programme (NBMP) (Bat Conservation Trust, 2015, Barlow et al., 2015b).

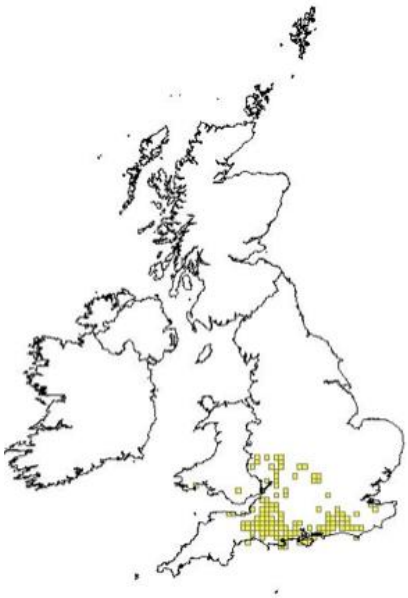
1. *C. barbastellus*

2. *E. serotinus*

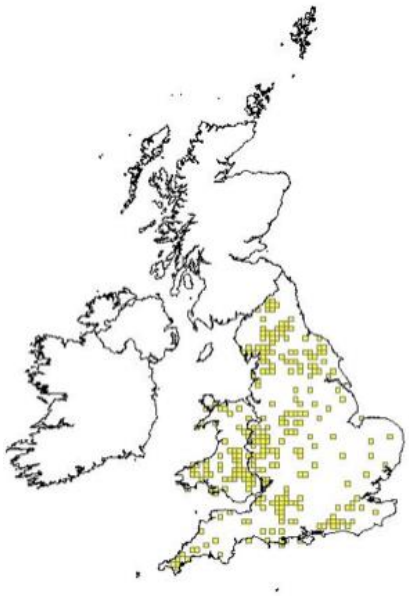


3. *M. alcathoe*

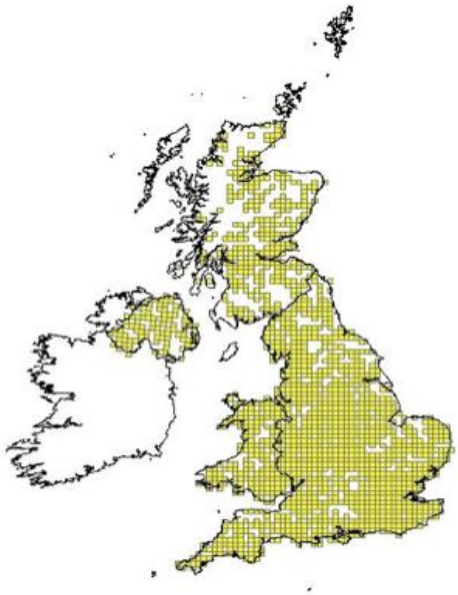
4. *M. bechsteinii*



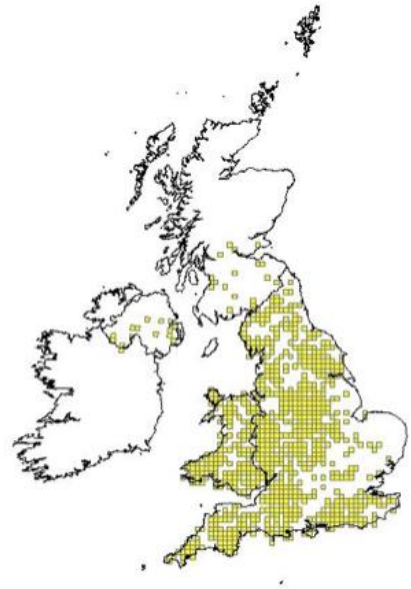
5. *M. brandtii*



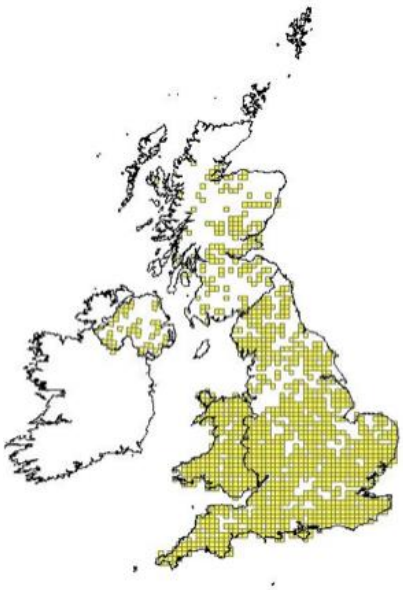
6. *M. daubentonii*



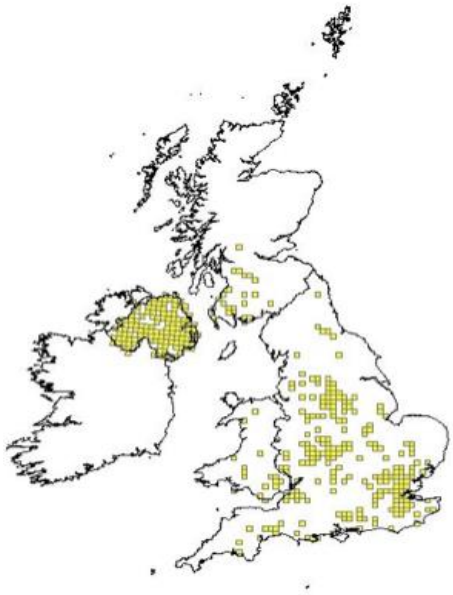
7. *M. mystacinus*



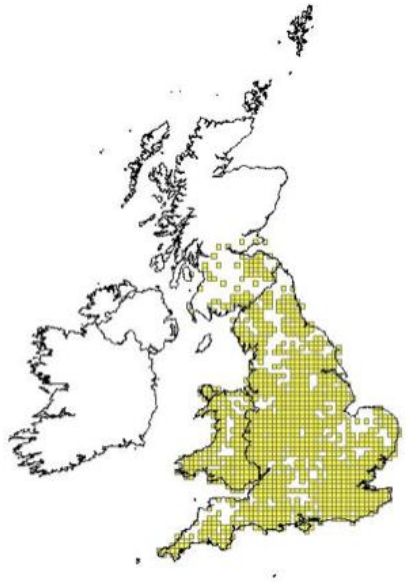
8. *M. nattereri*



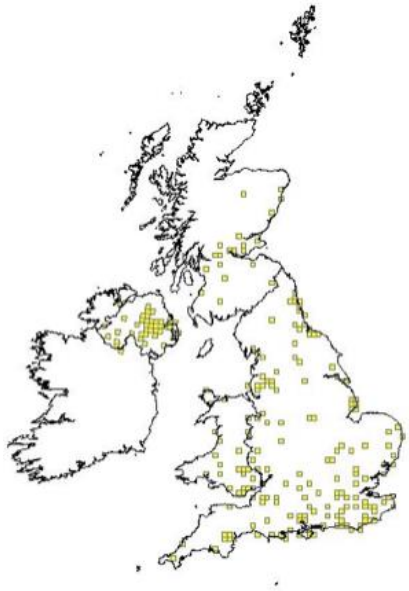
9. *N. leisleri*



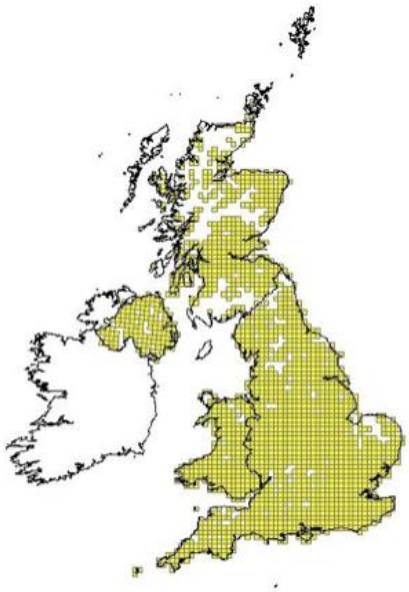
10. *N. noctula*



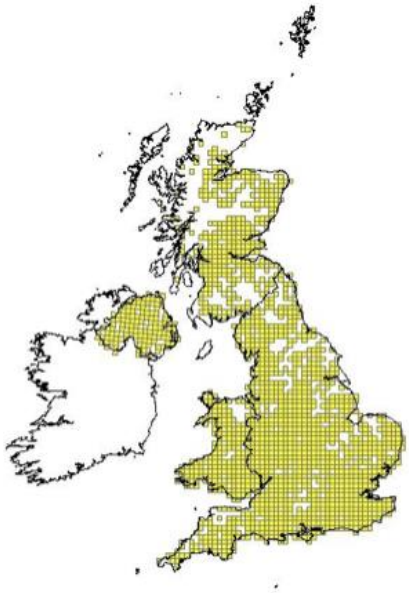
11. *P. nathusii*



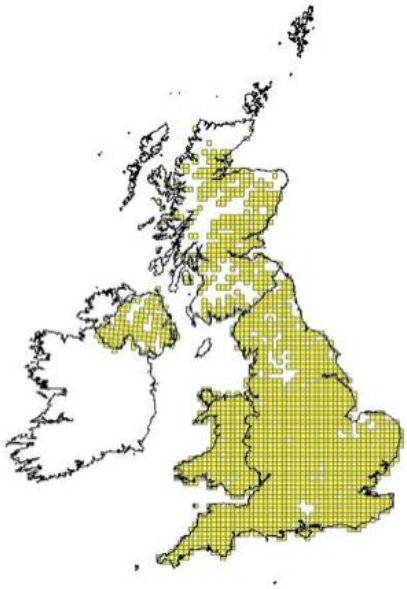
12. *P. pipistrellus*



13. *P. pygmaeus*

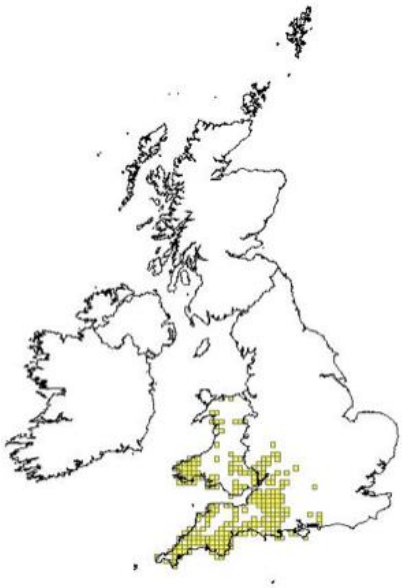


14. *P. auritus*

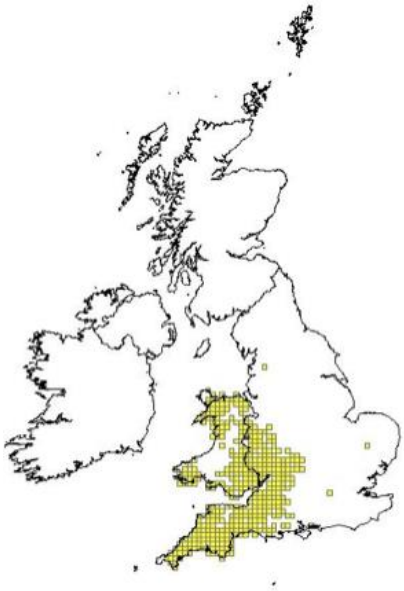


15. *P. austriacus*

16. *R. ferrumequinum*



17. *R. hipposideros*



Chapter 2:

C.2.1. Analysis of the literature on the diets of bats sampled in Great Britain

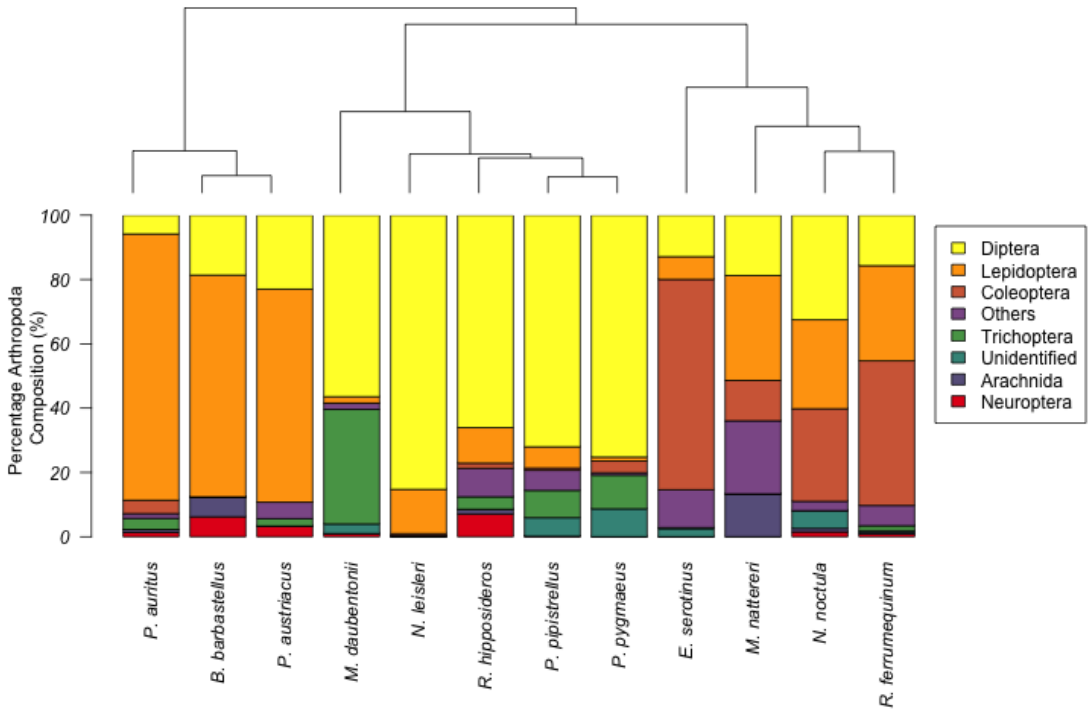


Figure C.2.1.1: The diet of each bat species from the literature sampled in Great Britain. The average diets of each of the species using all of the methods. Prey taxa have been grouped by order (where possible), class, phylum, or kingdom (where necessary). Diets sorted using complete-linkage clustering.

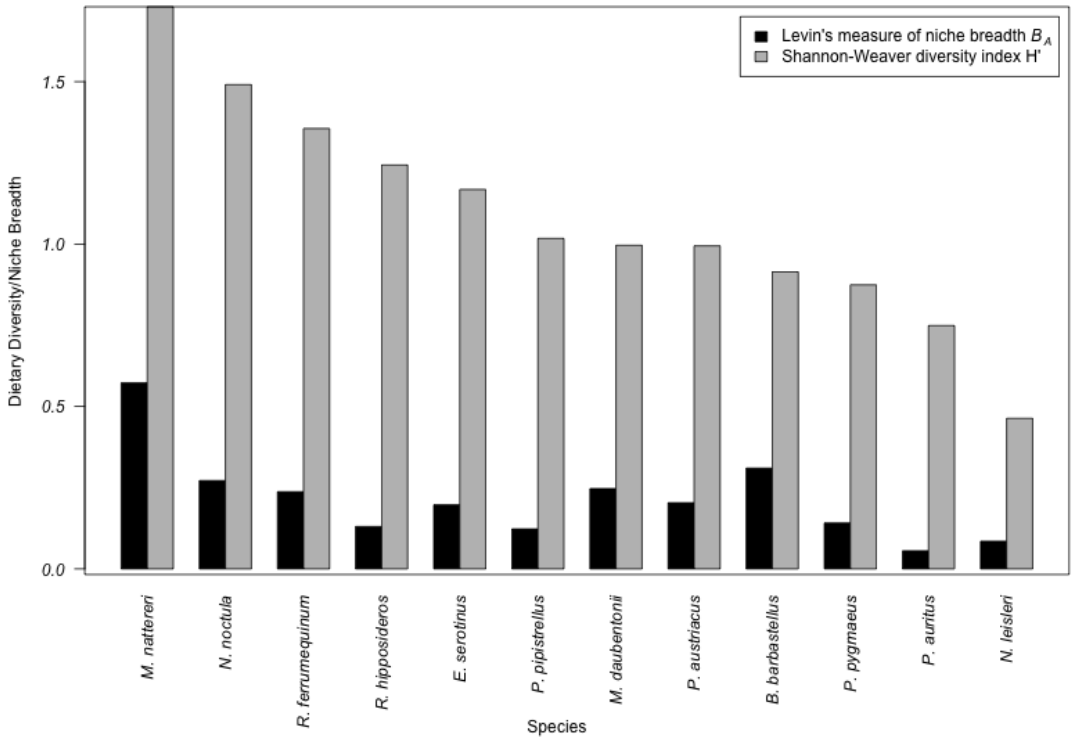


Figure C.2.1.2: Dietary diversity and niche breadth of each species. The dietary diversity calculated using Shannon-Weaver diversity index (H) (grey) and niche breadth calculated using Levins standardised index (B_A) (black).

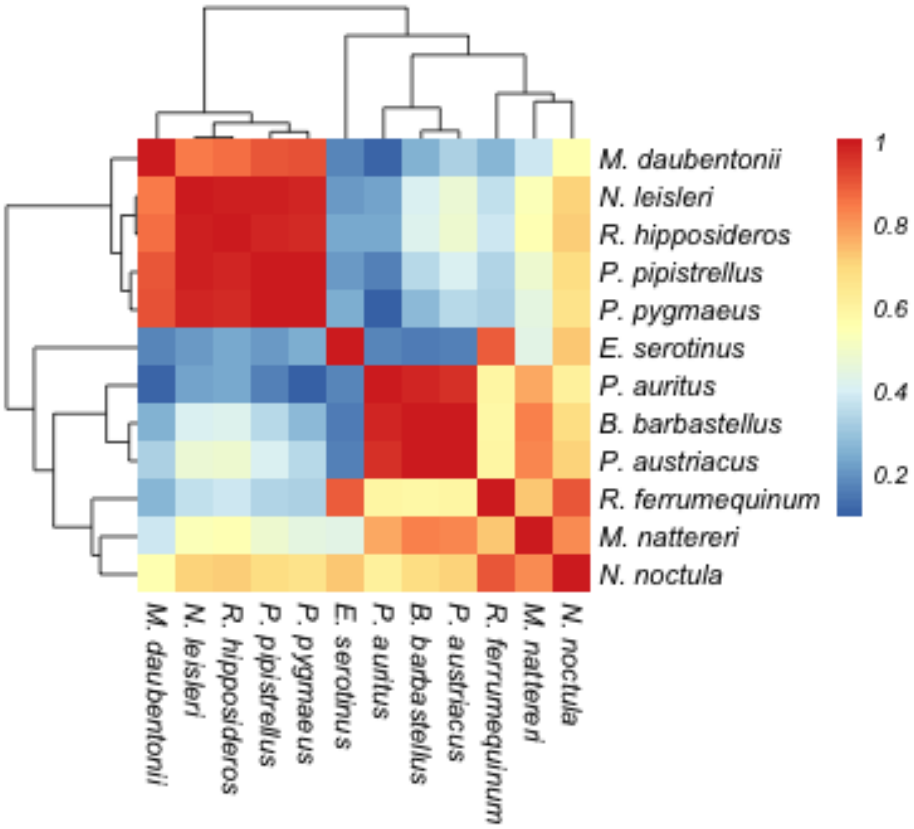


Figure C.2.1.3: The dietary overlap of each species. Calculated using Pianka's index of niche overlap. A value of 1 (red) suggests that the diets are identical, whilst a value of 0 (blue) indicates that there is no overlap. Species have been hierarchically clustered by niche overlap similarity.

Chapter 3:

C.3.1. Fragmentation study

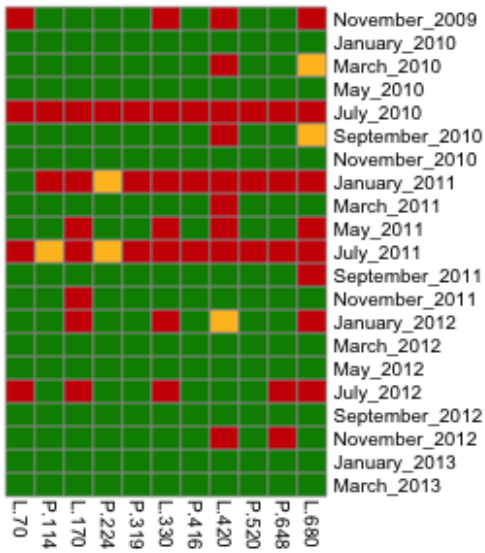


Figure C.3.1.1: PCR successes of samples. Ordered by date received (top to bottom) and size of amplicon (left to right). P= *Plecotus*, L= Lepidoptera. Red= no amplification, Yellow= faint band, Green= PCR success.

C.3.2. FastQC Plots

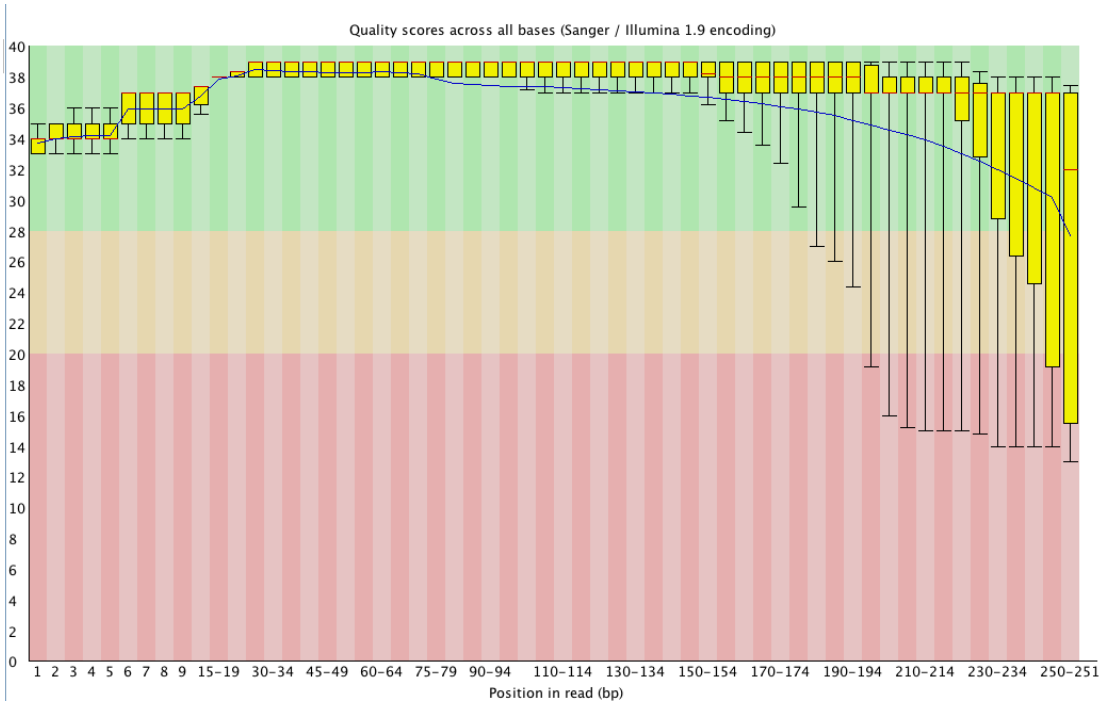


Figure C.3.2.1: FastQC plot showing the quality score across all bases for the first MiSeq Run. After running cut-adapt.

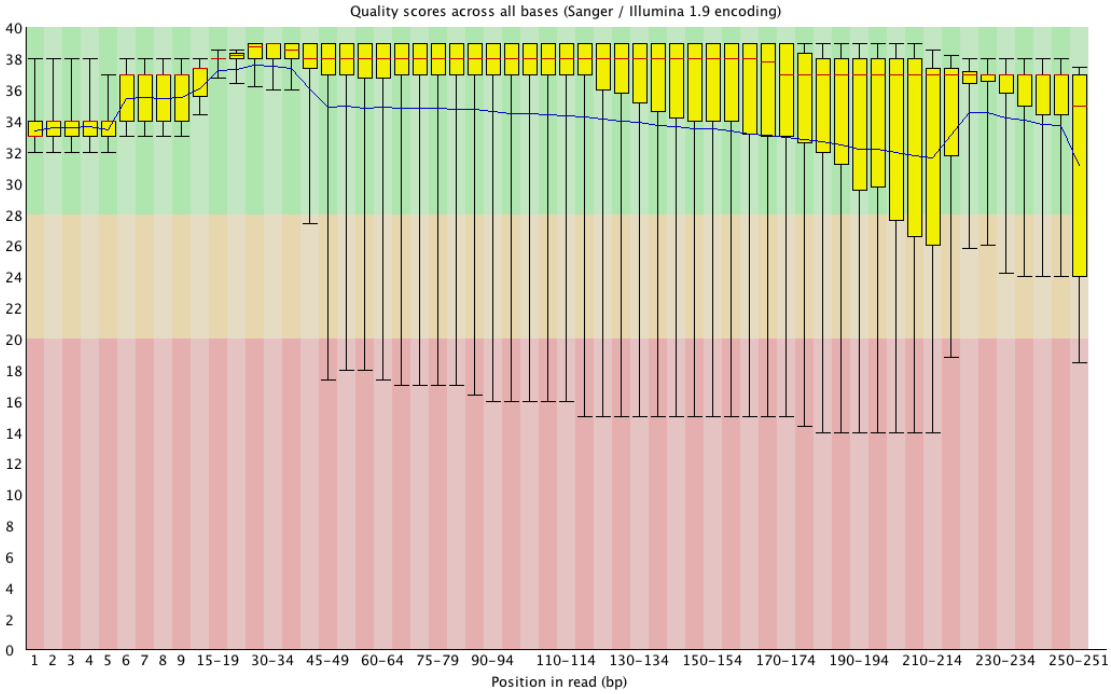


Figure C.3.2.2: FastQC plot showing the quality score across all bases for the second MiSeq Run. After running cut-adapt

C.3.3. Rarefaction plots

Figure C.3.3.1. shows the rarefaction plot for the Chiroptera subset at genus level. *N. leisleri* has the highest number of leaves, which is likely due to the misassignment of *N. leisleri* Chiroptera reads due to the limited representation of *N. leisleri* data in the NCBI nt database (table 3.10). Most of the rarefaction curves plateau quickly, except *C. barbastellus* and *M. brandtii*.

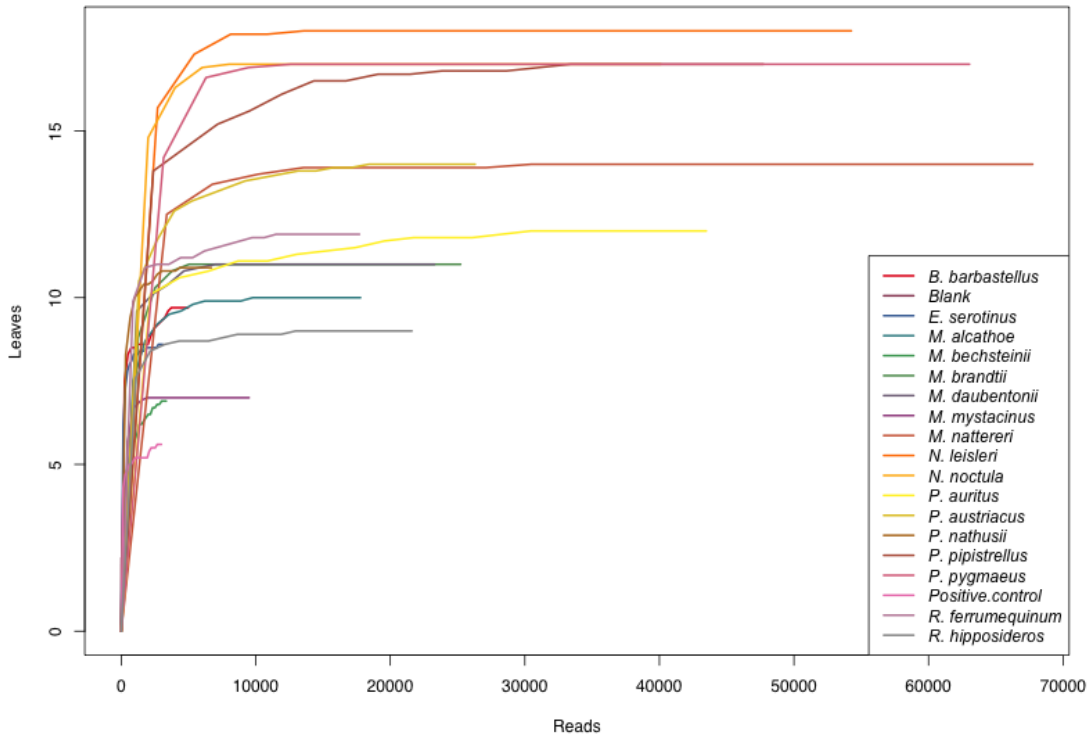


Figure C.3.3.1: Rarefaction plot for the Chiroptera subset at genus level. Including controls. Code in appendix D.

Figure C.3.3.2 shows the rarefaction plot for the Arthropoda subset of the data, collapsed at order level. Again, most curves plateau quickly, suggesting that sampling effort has been sufficient to find most guano arthropod taxa. *R. ferrumequinum,* which has the largest numbers of leaves (when collapsed to order level), and is the slowest rarefaction

curve to plateau, suggesting that a greater sampling effort is required to identify all Arthropoda, however, most have likely been identified.

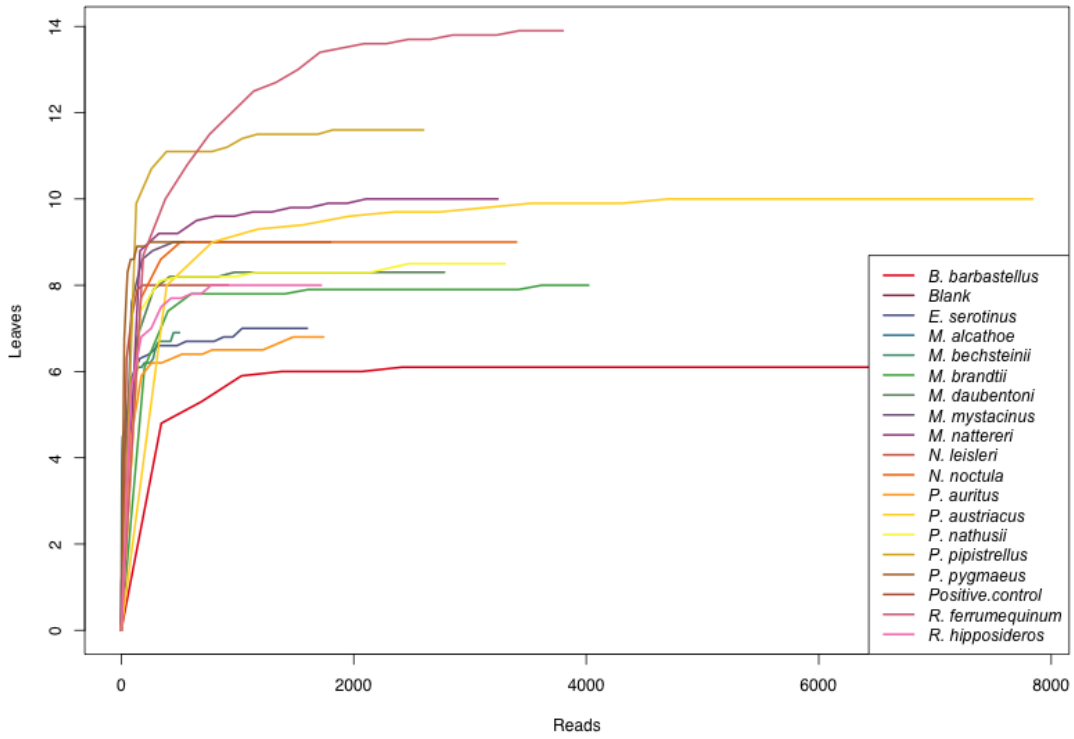


Figure C.3.3.2. Rarefaction plot for the Arthropoda subset at order level. Including controls. Code in D.3.10.

C.3.4. Species level Data assignments

Figure C.3.4 shows the assignment of the metagenome Arthropoda data assigned to species level and above. Colours are grouped by Order.

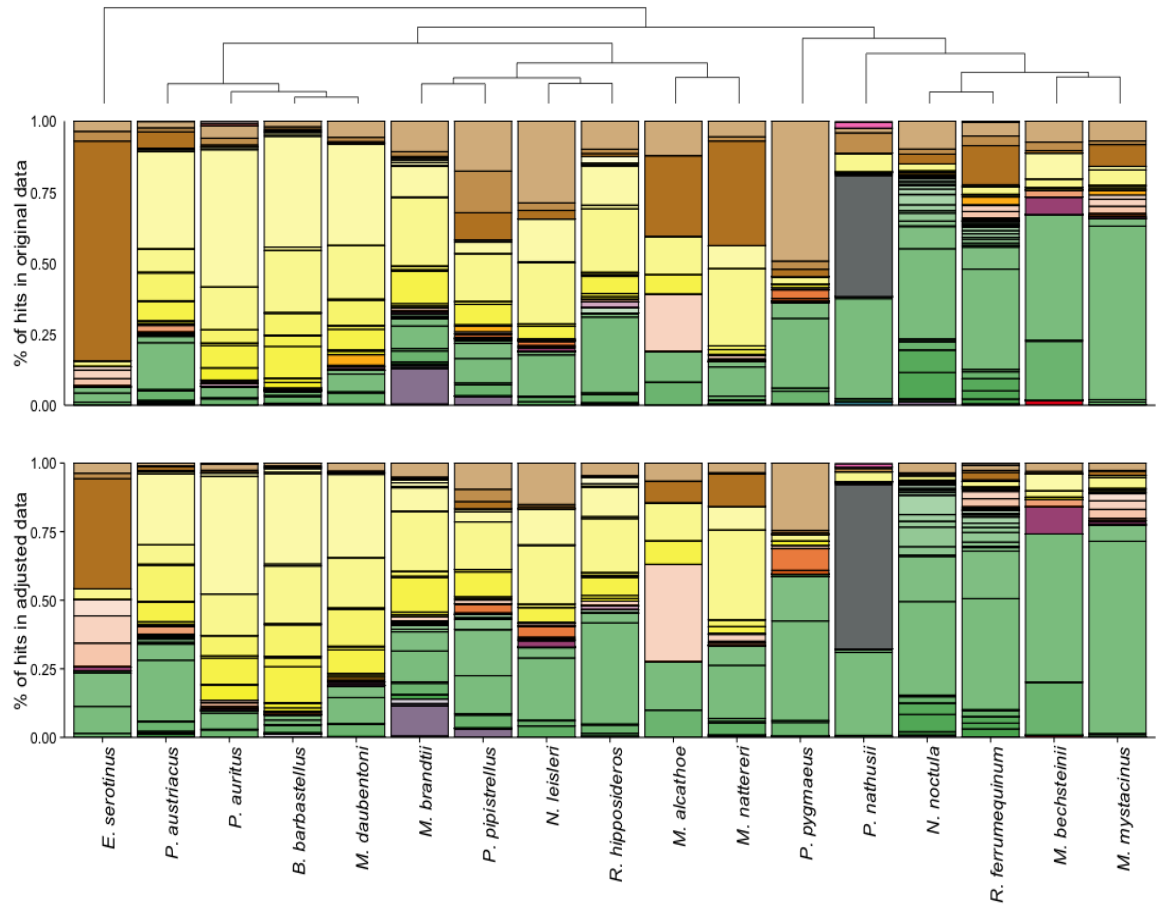


Figure C.3.4.1: The diet of each bat species as assigned at species level before and after PIA.

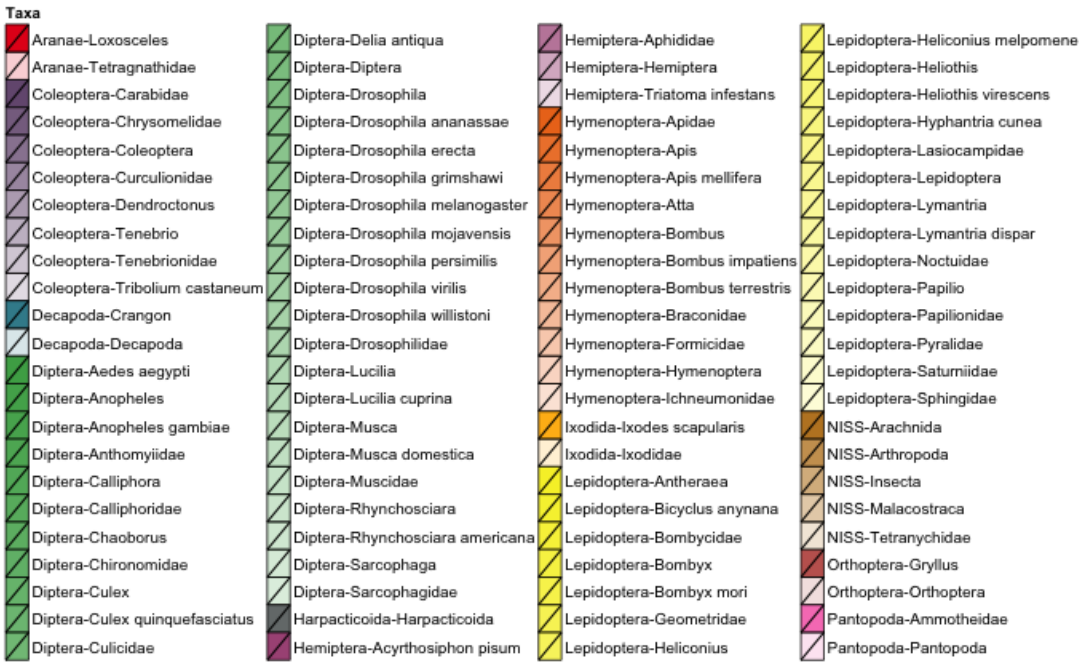


Figure C.3.4.2: Legend.

Chapter 4:

C.4.1. FastQC Plots

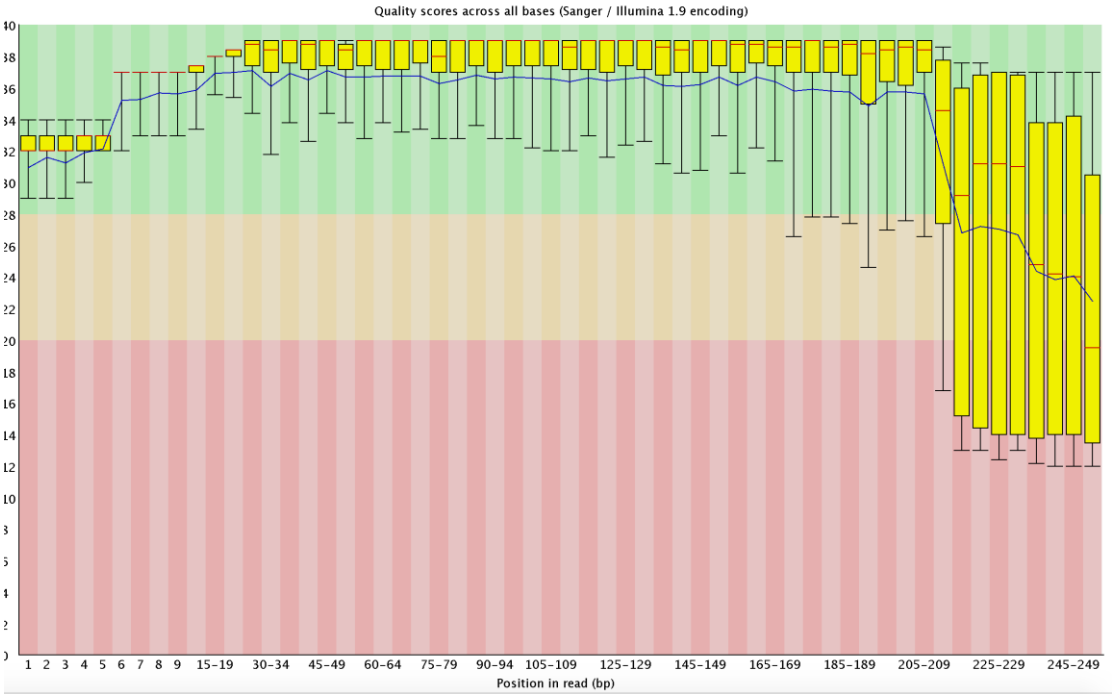


Figure C.4.1: FastQC plot showing the quality score across all bases. After running cut-adapt.

Chapter 5:

No Figures

Chapter 6:

No Figures

Chapter 7:

No Figures

Appendix D: Code

All code is stored in the digital appendix and is as follows:

D.1. Perl Code

D.1.1.PIA.pl

D.1.1.1. Helpfile_PIA.txt

D.1.2. PIA_to_GSA.pl

D.1.3. GSA.pl

D.1.3.1. Helpfile_GSA.txt

D.1.4. Modules

D.1.4.1. DebugTools.pm

D.1.4.2. FastaCheck.pm

D.1.4.3. FileChecks.pm

D.1.4.4. FileManipulations.pm

D.1.4.5. FileMerge.pm

D.1.4.6. GenomeSize.pm

D.1.4.7. TreeOfLife.pm

D.1.5. Check_primer_GC.pl

D.1.6. Primer_reverse_complement.pl

D.2. R Code

D.2.1. Methods_analysis_figure.r

D.2.2. Stacked_barchart.r

D.2.3. Dietary_diversity_and_breadth.r

D.2.4. Heatmap.r

D.2.5. Map.r

D.2.6. Sequence_length_distribution.r

D.2.7. Scatterplot.r

D.2.8. Pie_chart.r

D.2.9. Rarefaction.r

D.2.10. Barplot.r

D.2.11. PCA.r

Appendix E: Tables

Chapter 1:

No tables

Chapter 2:

E.2.1. References used to characterise diets.

Reference	Year of publication	Species studied	Country of sample collection	Reference
Ahmim & Moali	2013	*R. ferrumequinum, R. hipposideros*	Algeria	(Ahmim and Moali, 2013)
Andreas	2010	*P. auritus, P. austriacus*	Czech Republic	(Andreas, 2010)
Andreas et al.	2013	*R. ferrumequinum, R. hipposideros*	Slovakia	(Andreas et al., 2013)
Andreas et al.	2012	*B. barbastellus*	Czech Republic	(Andreas et al., 2012b)
Andreas et al.	2012	*M. bechsteinii, M. nattereri , P. auritus*	Central Europe	(Andreas et al., 2012a)
Arlettaz et al.	2000	*P. pipistrellus, R. hipposideros*	Switzerland	(Arlettaz et al., 2000)
ASHG	1994	*R. ferrumequinum*	Switzerland	(Beck et al., 1994)
Ashrafi et al.	2011	*P. auritus, P. austriacus*	Switzerland	(Ashrafi et al., 2011)
Barlow	1997	*P. pipistrellus, P. pygmaeus*	Britain	(Barlow, 1997)
Barta	1975	*P. auritus*	Slovak Republic	(Bárta, 1975)
Bartonička et al.	2008	*P. pygmaeus*	Czech Republic	(Bartonička et al., 2008)
Bauerová	1982	*P. austriacus*	Czech Republic	(Bauerova, 1982)
Bauerova & Cerveny	1986	*M. nattereri*	Czech Republic	(Bauerová and Cerveny, 1986)
Beck	1995	*B. barbastellus, E. serotinus, M. mystacinus, N. leisleri, N. noctula, P. auritus, P. austriacus, P.*	Switzerland	(Beck, 1995)

		nathusii, P. pipistrellus, R. ferrumequinum		
Beck	1991	*M. nattereri*	Switzerland	(Beck, 1991)
Beck et al.	1989	*R. hipposideros*	Switzerland	(Beck et al., 1989)
Bontadina et al.	2008	*R. hipposideros*	Switzerland	(Bontadina et al., 2008)
Boonman	1995	*P. auritus, R. ferrumequinum*	Netherlands, Belgium	(Boonman, 1995)
Buckhurst	1930	*P. auritus*	Britain	(Buckhurst, 1930)
Catto et al.	1994	*E. serotinus*	England	(Catto et al., 1994)
Chung et al.	2015	*E. serotinus*	Korea	(Chung et al., 2015)
Danko et al.	2010	*M. alcathoe*	Slovakia	(Danko et al., 2010)
Duvergé	1997	*R. ferrumequinum*	Britain	(Vaughan, 1997)
Feldman et al.	2000	*P. austriacus, R. hipposideros*	Israel	(Feldman et al., 2000)
Flanders & Jones	2009	*R. ferrumequinum*	Britain	(Flanders and Jones, 2009)
Flavin et al.	2001	*M. daubentonii*	Ireland	(Flavin et al., 2001)
Gajdosik & Gaisler	2004	*E. serotinus*	Czech Republic	(Gajdosik and Gaisler, 2004)
Gerber et al.	1994	*E. serotinus*	Switzerland	(Gerber et al., 1994)
Gloor et al.	1989	*N. noctula*	Switzerland	(Gloor et al., 1989)
Hanson	1950	*P. auritus*	Sweden	(Vaughan, 1997)
Heinicke & Krauß	1978	*P. auritus*	Germany	(Heinicke and Krauß, 1978)

Hoare	1991	*P. pipistrellus*	England	(Hoare, 1991)
Hollyfield	1993	*M. mystacinus, P. auritus, R. hipposideros*	Ireland	(Vaughan, 1997)
Hope et al.	2014	*M. nattereri*	England	(Hope et al., 2014)
Jin et al.	2005	*R. ferrumequinum*	China	(Jin et al., 2005)
Jing et al.	2010	*R. ferrumequinum*	China	(Wang Jing, 2010)
Jones	1995	*N. noctula*	Britain	(Jones, 1995)
Jones	1990	*R. ferrumequinum*	China	(Jones, 1990)
Kaňuch et al.	2005	*N. noctula, N. leisleri*	Slovakia, Czech Republic	(Kaňuch et al., 2005a)
Kaňuch et al.	2005	*N. leisleri*	Slovakia	(Kaňuch et al., 2005b)
Kervyn & Libois	2008	*E. serotinus*	Belgium	(Kervyn and Libois, 2008)
Krauss	1978	*P. auritus*	Germany	(Krauss, 1978)
Kruger et al.	2013	*M. daubentonii*	Germany	(Krüger et al., 2013a)
Kruger et al.	2013	*P. nathusii*	Latvia	(Krüger et al., 2013b)
Kruger et al.	2012	*M. daubentonii*	Germany	(Krüger et al., 2012)
Leishman	1983	*R. ferrumequinum, R. hipposideros*	Britain	(Vaughan, 1997)
Lino et al.	2014	*R. hipposideros*	Portugal	(Lino et al., 2014)
Lucan et al.	2009	*M. alcathoe*	Czech Republic	(Lucan et al., 2009)

Ma et al.	2008	*R. ferrumequinum*	China	(Ma et al., 2008)
Mackenzie & Oxford	1995	*N. noctula*	Britain	(Mackenzie and Oxford, 1995)
Manwaring-Banes	1939	*P. auritus*	Britain	(Manwaring, 1939)
McAney & Fairley	1989	*R. hipposideros*	Ireland	(McAney and Fairley, 1989)
McAney et al.	1991	*E. serotinus*	Czech Republic	(McAney, 1991)
MIkula & Čmoková	2012	*E. serotinus*	Czech Republic	(Mikula and Čmoková, 2012)
Nissen et al.	2013	*M. daubentonii*	Germany	(Nissen et al., 2013)
Oldfield	1990	*P. auritus*	Sweden	(Vaughan, 1997)
Pir	1994	*R. ferrumequinum*	Luxembourg	(Vaughan, 1997)
Poulton	1929	*R. ferrumequinum*	Britain	(Poulton, 1929)
Ransome	1996	*R. ferrumequinum*	Britain	(Vaughan, 1997)
Razgour et al.	2011	*P. auritus, P. austriacus*	Britain	(Razgour et al., 2011)
Robertson	1988	*R. ferrumequinum*	Britain	(Vaughan, 1997)
Robinson	1990	*P. auritus*	Britain	(Robinson, 1990)
Robinson & Stebbings	1993	*E. serotinus*	Britain	(Robinson and Stebbings, 1993)
Rostovskaya et al.	2000	*P. auritus*	Central Russia	(Rostovskaya et al., 2000)
Roswag et al.	2015	*M. bechsteinii, M. nattereri, P. auritus*	Central Germany	(Roswag et al., 2015)
Rydell	1989	*P. auritus*	Sweden	(Rydell, 1989b)

Rydell et al.	1996	*B. barbastellus*	Germany, Sweden	(Rydell et al., 1996)
Seimers & Swift	2006	*M. bechsteinii, M. nattereri*	Germany	(Siemers and Swift, 2006)
Sheil et al.	1991	*M. nattereri, P. auritus*	Ireland	(Shiel et al., 1991)
Shiel et al.	1998	*N. leisleri*	Ireland, England	(Shiel et al., 1998)
Sierro & Arlettaz	1997	*B. barbastellus*	Switzerland, Asia	(Sierro and Arlettaz, 1997)
Smirnov & Vekhnik	2014	*E. serotinus, M. brandtii, M. daubentonii, M. mystacinus, M. nattereri, N. leisleri, N. noctula, P. auritus, P. nathusii, P. pipistrellus, M. daubentonii*	Russia	(Smirnov and Vekhnik, 2014)
Sologor	1980	*E. serotinus*	Ukraine	(Sologor, 1980)
Sullivan et al.	1993	*M. daubentonii, N. leisleri, P. pipistrellus*	Ireland	(Sullivan et al., 1993)
Swift & Racey	1983	*M. daubentonii, P. auritus*	Scotland	(Swift and Racey, 1983)
Swift et al.	1985	*P. pipistrellus*	Scotland	(Swift et al., 1985)
Taake	1993	*M. bechsteinii, M. brandtii, M. daubentonii, M. mystacinus, M. nattereri , P. auritus*	Germany	(Taake, 1993)
Thompson	1982	*P. auritus*	England	(Thompson, 1982)
Vesterinen et	2013	*M. daubentonii*	Finland	(Vesterinen et al.,

al.				2013)
Walhovd & Hoegh-Gildberg	1984	*P. auritus*	Denmark	(Vaughan, 1997)
Waters et al.	1999	*N. leisleri*	Britain	(Waters et al., 1999)
Waters et al.	1995	*N. leisleri*	Britain	(Waters et al., 1995)
Whitaker	1994	*M. nattereri , P. austriacus*	Israel	(Vaughan, 1997)
Whitaker & Karataş	2009	*B. barbastellus, E. serotinus, M. brandtii, M. mystacinus, M. nattereri , P. auritus, P. austriacus, P. pipistrellus, P. pygmaeus, R. ferrumequinum, R. hipposideros*	Turkey	(Whitaker Jr and Karatas, 2009)
Williams et al.	2011	*R. hipposideros*	Britain	(Williams et al., 2010)
Wolz	1993	*M. bechsteinii*	Germany	(Vaughan, 1997)
Zeale et al.	2011	*B. barbastellus, M. nattereri, P. pipistrellus*	Britain	(Zeale et al., 2011)
Zukal & Gajdošík	2012	*E. serotinus*	Czech Republic	(Zukal and Gajdošík, 2012)

Chapter 3:

E.3.1. NCBI nt database coverage for each of the Great British bat species. As of 5th February 2016 (N.C.B.I. Resource Coordinators, 2013).

Species	**Reads from the whole genome**	**Reads from COI**
B. barbastellus	149	43

E. serotinus	602	29
M. alcathoe	204	5
M. bechsteinii	125	9
M. brandtii	386614	70
M. daubentonii	464	47
M. mystacinus	245	28
M. nattereri	269	28
N. leisleri	127	12
N. noctula	152	6
P. auritus	197	16
P. austriacus	274	2
P. nathusii	57	14
P. pipistrellus	610	36
P. pygmaeus	289	13
R. ferrumequinum	70945	10
R. hipposideros	1035	6

E.3.2. The Chiroptera genera represented in the dataset, and their representation in the Genbank nucleotide (nt) database.

Genera	**Total Genbank nt entries**	**COI Genbank entries**
Antrozous	209	7
Artibeus	5207	2890
Barbastella	185	45
Carollia	2730	1563
Chalinolobus	62	8
Eptesicus	201016	246
Glauconycteris	41	8
Kerivoula	576	342
Lasiurus	2283	492
Murina	6163	491
Myotis	976510	1308
Neoromicia	437	64
Nyctalus	427	19
Pipistrellus	2365	430
Plecotus	1318	64
Pteropus	253453	103
Rhinolophus	82233	813
Rhogeessa	316	32

Scotophilus	514	55
Tylonycteris	193	38

E.3.3. The Arthropoda species represented in the dataset, and their representation in the Genbank nucleotide (nt) database. Arthropoda collated to order level as of 15th February 2016.

Arthropoda	**Genbank nt representation**
Astigmata	10213
Ixodida	941661
Mesostigmata	120740
Araneae	680473
Opiliones	9548
Pantopoda	2198
Decapoda	958247
Harpacticoida	76552
Lepidoptera	2307559
Trichoptera	197348
Coleoptera	896319
Diptera	3057765
Hymenoptera	3475363
Neuroptera	71873
Siphonaptera	93434
Dermaptera	27025
Orthoptera	173961
Hemiptera	1679609
Ephemeroptera	21658

E.3.4. Sample details and result summaries-digital appendix

E.3.5. Genome size spreadsheet-digital appendix

Chapter 4:

E.4.1. Sample details and result summaries- digital appendix

Chapter 5:

No tables

Chapter 6:

E.6.1. Sample details and result summaries- digital appendix

Chapter 7:

No tables

Appendix F: Publications

Publications not directly using data from this thesis

Kistler, L., Smith, O., Ware, R., Momber, G., Bates, R., Garwood, P., Fitch, S., Pallen, M., Gaffney, V., Allaby, R. (2015). Thermal age, cytosine deamination and the veracity of 8,000-year-old wheat DNA from sediments. *bioRxive.* 032060.

Allaby, R., Gutaker, R., Clarke, A., Pearson, N., Ware, R., Palmer, S., Kitchen, J., Smith, O. (2015) Using archaeogenomic and computational approaches to unravel the history of local adaptation in crops. *Philosophical Transactions B*. 370: 20130377

Allaby, R., Kistler, L., Gutaker, R., Ware, R., Kitchen, J., Smith, O., Clarke, A. (2015) Archeogenomic insights into the adaptation of plants to the human environment: pushing plant-hominin co-evolution back to the Pliocene. *Journal of Human Evolution.* 79 (150-157)

Bibliography

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V. & Polz, M. F. 2005. PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology,* **71**, (12): 8966-8969.

Adams, R. A. & Thibault, K. M. 2006. Temporal resource partitioning by bats at water holes. *Journal of Zoology,* **270**, (3): 466-472.

Aguirre, L. F., Herrel, A., Van Damme, R. & Matthysen, E. 2002. Ecomorphological analysis of trophic niche partitioning in a tropical savannah bat community. *Proceedings of the Royal Society B: Biological Sciences,* **269**, (1497): 1271-1278.

Ahmim, M. & Moali, A. 2013. The diet of four species of horseshoe bat (Chiroptera: Rhinolophidae) in a mountainous region of Algeria: evidence for gleaning. *Hystrix, the Italian Journal of Mammalogy,* **24**, (2): 4.

Alexander, J. B., Newton, J. & Crow, G. A. 1991. Distribution of Oriental and German cockroaches, *Blatta orientalis* and *Blattella germanica* (Dictyoptera), in the United Kingdom. *Medical and Veterinary Entomology,* **5**, (4): 395-402.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. 1990. Basic local alignment search tool. *Journal of Molecular Biology,* **215**, (3): 403-410.

Andreas, M. 2010. Study on bat diet in the Czech Republic and Slovakia. *A Tribute to Bats, Prague*: 261-264.

Andreas, M., Reiter, A. & Benda, P. 2012a. Dietary composition, resource partitioning and trophic niche overlap in three forest foliage-gleaning bats in Central Europe. *Acta Chiropterologica,* **14**, (2): 335-345.

Andreas, M., Reiter, A. & Benda, P. 2012b. Prey selection and seasonal diet changes in the western barbastelle bat (*Barbastella barbastellus*). *Acta Chiropterologica,* **14**, (1): 81-92.

Andreas, M., Reiter, A., Cepáková, E. & Uhrin, M. 2013. Body size as an important factor determining trophic niche partitioning in three syntopic rhinolophid bat species. *Biologia,* **68**, (1): 170-175.

Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data. *Reference Source*.

Aperis, G., Fuchs, B. B., Anderson, C. A., Warner, J. E., Calderwood, S. B. & Mylonakis, E. 2007. *Galleria mellonella* as a model host to study infection by the Francisella tularensis live vaccine strain. *Microbes and Infection,* **9**, (6): 729-734.

Arata, A. A., Vaughn, J. B., Newell, K. W., Barth, R. A. & Gracian, M. 1968. Salmonella and Shigella infections in bats in selected areas of Colombia. *American Journal of Tropical Medicine and Hygiene,* **17**, (1): 92-5.

Arlettaz, R. 1999. Habitat selection as a major resource partitioning mechanism between the two sympatric sibling bat species *Myotis myotis* and *Myotis blythii*. *Journal of Animal Ecology,* **68**, (3): 460-471.

Arlettaz, R., Godat, S. & Meyer, H. 2000. Competition for food by expanding pipistrelle bat populations (*Pipistrellus pipistrellus*) might contribute to the decline of lesser horseshoe bats (*Rhinolophus hipposideros*). *Biological Conservation,* **93**, (1): 55-60.

Arlettaz, R., Perrin, N. & Hausser, J. 1997. Trophic resource partitioning and competition between the two sibling bat species *Myotis myotis* and *Myotis blythii*. *Journal of Animal Ecology,* **66**, (6): 897-911.

Ashrafi, S., Beck, A., Rutishauser, M., Arlettaz, R. & Bontadina, F. 2011. Trophic niche partitioning of cryptic species of long-eared bats in Switzerland: implications for conservation. *European Journal of Wildlife Research,* **57**, (4): 843-849.

Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A. & Gordon, J. I. 2005. Host-bacterial mutualism in the human intestine. *Science,* **307**, (5717): 1915-1920.

Barber, J. R. & Conner, W. E. 2007. Acoustic mimicry in a predator–prey interaction. *Proceedings of the National Academy of Sciences,* **104**, (22): 9331-9334.

Barber, J. R., Leavell, B. C., Keener, A. L., Breinholt, J. W., Chadwell, B. A., Mcclure, C. J., Hill, G. M. & Kawahara, A. Y. 2015. Moth tails divert bat attack: evolution of

acoustic deflection. *Proceedings of the National Academy of Sciences,* **112**, (9): 2812-2816.
Barclay, R. M., Dolan, M.-A. & Dyck, A. 1991. The digestive efficiency of insectivorous bats. *Canadian Journal of Zoology,* **69**, (7): 1853-1856.
Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y. & Van Wezel, G. P. 2016. Taxonomy, physiology, and natural products of Actinobacteria. *Microbiology and Molecular Biology Reviews,* **80**, (1): 1-43.
Barlow, A., Worledge, L., Miller, H., Drees, K., Wright, P., Foster, J. T., Sobek, C., Borman, A. & Fraser, M. 2015a. First confirmation of *Pseudogymnoascus destructans* in British bats and hibernacula. *Veterinary Record,* **177**, (3): 73.
Barlow, K., Briggs, P., Haysom, K., Hutson, A., Lechiara, N., Racey, P., Walsh, A. & Langton, S. 2015b. Citizen science reveals trends in bat populations: the national bat monitoring programme in Great Britain. *Biological Conservation,* **182**: 14-26.
Barlow, K. & Jones, G. 1996. *Pipistrellus nathusii* (Chiroptera: Vespertilionidae) in Britain in the mating season. *Journal of Zoology,* **240**, (4): 767-773.
Barlow, K. E. 1997. The diets of two phonic types of the bat *Pipistrellus pipistrellus* in Britain. *Journal of Zoology,* **243**, (3): 597-609.
Barnard, P. C. 1999. Identifying British insects and arachnids: an annotated bibliography of key works*, Cambridge University Press.*
Barratt, E., Deaville, R., Burland, T., Bruford, M., Jones, G., Racey, P. & Wayne, R. 1997a. DNA answers the call of pipistrelle bat species. *Nature,* **387**, (6629): 138.
Barratt, E. M., Deaville, R., Burland, T. M., Bruford, M. W., Jones, G., Racey, P. A. & Wayne, R. K. 1997b. DNA answers the call of pipistrelle bat species. *Nature,* **387**, (6629): 138.
Bárta, Z. 1975. K potravě netopýra ušatého (*Plecotus auritus* L.). *Lynx, n. s,* **17**: 5-6.
Barton, H. A., Giarrizzo, J. G., Suarez, P., Robertson, C. E., Broering, M. J., Banks, E. D., Vaishampayan, P. A. & Venkateswaran, K. 2014. Microbial diversity in a Venezuelan orthoquartzite cave is dominated by the Chloroflexi (Class Ktedonobacterales) and Thaumarchaeota Group I. 1c. *Frontiers in Microbiology,* **5**.
Bartonička, T., Řehák, Z. & Andreas, M. 2008. Diet composition and foraging activity of *Pipistrellus pygmaeus* in a floodplain forest. *Biologia,* **63**, (2): 266-272.
Bat Conservation Trust 2009. The national bat monitoring programme. *Annual Report*.
Bat Conservation Trust. 2010. *Rarities and vagrants* [Online]. Bat Conservation Trust. Available: http://www.bats.org.uk/data/files/Species_Info_sheets/raritiesvagrants.pdf [Accessed 30 November 2015].
Bat Conservation Trust. 2013. *The bat sound library* [Online]. Available: http://www.bats.org.uk/pages/bat_sound_library_additional_notes.html [Accessed 15 December 2015].
Bat Conservation Trust 2014 Bat droppings: foundation course for ecological consultants. Bat Conservation Trust, London.
Bat Conservation Trust 2015. National bat monitoring programme. Annual report 2014. Bat Conservation Trust, London.
Bates, H. W. 1981. Contributions to an insect fauna of the Amazon valley (Lepidoptera: Heliconidae). *Biological Journal of the Linnean Society,* **16**, (1): 41-54.
Bates, P. J., Nwe, T., Swe, K. M. & Bu, S. S. H. 2001. Further new records of bats from Myanmar (Burma), including *Craseonycteris thonglongyai* (Hill, 1974) (Chiroptera: Craseonycteridae). *Acta Chiropterologica,* **3**, (1): 33-41.
Battersby, J. 2005. UK mammals: Species status and population trends. *JNCC/Tracking Mammals Partnership.*
Bauerova, Z. 1982. Contribution to the trophic ecology of the grey long-eared bat, *Plecotus austriacus*. *Folia Zoologica,* **31**: 113-122.
Bauerová, Z. & Cerveny, J. 1986. Towards an understanding of the trophic ecology of *Myotis nattereri*. *Folia Zoologica,* **35**, (1): 55-61.

Bearhop, S., Adams, C. E., Waldron, S., Fuller, R. A. & Macleod, H. 2004. Determining trophic niche width: a novel approach using stable isotope analysis. *Journal of Animal Ecology,* **73**, (5): 1007-1012.
Beck, A. 1991. Nahrungsuntersuchungen bei der Fransenfledermaus, *Myotis nattereri* (Kuhl, 1818). *Myotis,* **29**: 67-70.
Beck, A. 1995. Fecal analyses of European bat species. *Myotis,* **32**, (33): 109-119.
Beck, A., Bontadina, F., Gloor, S., Hotz, T., Lutz, M. & Muehlethaler, E. 1994. Jagdhabitatwahl und nächtliche Aufenthaltsgebiete der Grossen Hufeisennase (*Rhinolophus ferrumequinum*) im Raum Castrich/GR. *Arbeitsgruppe zum Schutz der Hufeisennasen Graubündens ASGH, Encarden,* **51**: 7152.
Beck, A., Stutz, H. & Ziswiler, V. 1989. Nutritional habits of the lesser horseshoe bat *Rhinolophus hipposideros* (Bechstein, 1800)(Mammalia, Chiroptera). *Revue Suisse de Zoologie,* **96**, (3): 643-650.
Benda, P., Hulva, P., Andreas, M. & Uhrin, M. 2003. Notes on the distribution of *Pipistrellus pipistrellus* complex in the Eastern Mediterranean: First records of *P. pipistrellus* for Syria and of *P. pygmaeus* for Turkey. *Vespertilio,* **7**: 87-95.
Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. 2012. GenBank. *Nucleic Acids Research*: gks1195.
Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A. & Wheeler, D. L. 2000. GenBank. *Nucleic Acids Research,* **28**, (1): 15-18.
Best, K., Oakes, T., Heather, J. M., Shawe-Taylor, J. & Chain, B. 2015. Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding. *Scientific Reports,* **5**: 14629.
Bethea, D. M., Carlson, J. K., Buckel, J. A. & Satterwhite, M. 2006. Ontogenetic and site-related trends in the diet of the Atlantic sharpnose shark *Rhizoprionodon terraenovae* from the northeast Gulf of Mexico. *Bulletin of Marine Science,* **78**, (2): 287-307.
Bird, M. I., Boobyer, E. M., Bryant, C., Lewis, H. A., Paz, V. & Stephens, W. E. 2007. A long record of environmental change from bat guano deposits in Makangit Cave, Palawan, Philippines. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh,* **98**, (01): 59-69.
Blackwood, J. C., Streicker, D. G., Altizer, S. & Rohani, P. 2013. Resolving the roles of immunity, pathogenesis, and immigration for rabies persistence in vampire bats. *Proceedings of the National Academy of Sciences,* **110**, (51): 20837-20842.
Bogdanowicz, W., Csada, R. D. & Fenton, M. B. 1997. Structure of noseleaf, echolocation, and foraging behavior in the Phyllostomidae (Chiroptera). *Journal of Mammalogy,* **78**, (3): 942-953.
Bogdanowicz, W., Fenton, M. B. & Daleszczyk, K. 1999. The relationships between echolocation calls, morphology and diet in insectivorous bats. *Journal of Zoology,* **247**, (3): 381-393.
Bohmann, K., Monadjem, A., Noer, C. L., Rasmussen, M., Zeale, M. R. K., Clare, E., Jones, G., Willerslev, E. & Gilbert, M. T. P. 2011. Molecular diet analysis of two African free-tailed bats (Molossidae) using high throughput sequencing. *PLoS ONE,* **6**, (6): e21441.
Bon, C., Berthonaud, V., Maksud, F., Labadie, K., Poulain, J., Artiguenave, F., Wincker, P., Aury, J.-M. & Elalouf, J.-M. 2012. Coprolites as a source of information on the genome and diet of the cave hyena. *Proceedings of the Royal Society of London B: Biological Sciences*: rspb20120358.
Bontadina, F., Schmied, S. F., Beck, A. & Arlettaz, R. 2008. Changes in prey abundance unlikely to explain the demography of a critically endangered Central European bat. *Journal of Applied Ecology,* **45**, (2): 641-648.
Boonman, A. M., Boonman, M., Bretschneider, F. & Van De Grind, W. A. 1998. Prey detection in trawling insectivorous bats: duckweed affects hunting behaviour in Daubenton's bat, *Myotis daubentonii*. *Behavioral Ecology and Sociobiology,* **44**, (2): 99-107.
Boonman, M. 1995. Voedselkeuze grootoorvleermuis, *Zoogdier.* 6, 19-22.
Boshamer, J. P. C. & Bekker, J. P. 2008. Nathusius' pipistrelles (*Pipistrellus nathusii*) and other species of bats on offshore platforms in the Dutch sector of the North Sea. *Lutra,* **51**, (1): 17-36

Boyles, J. G. & Storm, J. J. 2007. The perils of picky eating: dietary breadth is related to extinction risk in insectivorous bats. *PLoS ONE,* **2**, (7): e672.
Bradley, J. D. & Bradley, M. J. 2000. Checklist of Lepidoptera recorded from the British Isles*, DJ Bradley & MJ Bradley*.
Brakenhoff, R. H., Schoenmakers, J. & Lubsen, N. H. 1991. Chimeric cDNA clones: a novel PCR artifact. *Nucleic Acids Research,* **19**, (8): 1949.
Bray, J. R. & Curtis, J. T. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs,* **27**, (4): 325-349.
Briggs, A. W., Stenzel, U., Johnson, P. L., Green, R. E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M. T. & Lachmann, M. 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences,* **104**, (37): 14616-14621.
Britzke, E., Gillam, E. & Murray, K. 2013. Current state of understanding of ultrasonic detectors for the study of bat ecology. *Acta Theriologica*: 1-9.
Brzezinska, M. S., Jankiewicz, U., Burkowska, A. & Walczak, M. 2014. Chitinolytic microorganisms and their possible application in environmental protection. *Current Microbiology,* **68**, (1): 71-81.
Buckhurst, A. 1930. Moths destroyed by a long eared bat. *Entomologist,* **63**: 238.
Cameron, S. L. & Whiting, M. F. 2008. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*,(Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene,* **408**, (1): 112-123.
Cannon, R. J. 1998. The implications of predicted climate change for insect pests in the UK, with emphasis on non‐indigenous species. *Global Change Biology,* **4**, (7): 785-796.
Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K. & Gordon, J. I. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods,* **7**, (5): 335-336.
Carey, H. V., Walters, W. A. & Knight, R. 2013. Seasonal restructuring of the ground squirrel gut microbiota over the annual hibernation cycle. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology,* **304**, (1): R33-R42.
Casadevall, A. 2012. Fungi and the rise of mammals. *PLoS Pathogens,* **8**, (8): e1002808.
Catto, C., Hutson, A., Raccey, P. & Stephenson, P. 1996. Foraging behaviour and habitat use of the serotine bat (*Eptesicus serotinus*) in southern England. *Journal of Zoology,* **238**, (4): 623-633.
Catto, C., Hutson, A. & Racey, P. 1994. The diet of *Eptesicus serotinus* in southern England. *Folia Zoologica,* **43**, (4): 307-314.
Charbonnier, Y., Barbaro, L., Theillout, A. & Jactel, H. 2014. Numerical and functional responses of forest bats to a major insect pest in pine plantations. *PLoS ONE,* **9**, (10): e109488.
Che, J., Huang, D., Li, D., Ma, J. & Zhang, Y. 2010. DNA barcoding and the international barcode of life project in China. *Biodiversity Informatics,* **24**, (4).
Chesmore, D. 2004. Automated bioacoustic identification of species. *Anais da Academia Brasileira de Ciências,* **76**, (2): 436-440.
Chibucos, M. C., Crabtree, J., Nagaraj, S., Chaturvedi, S. & Chaturvedi, V. 2013. Draft genome sequences of human pathogenic fungus *Geomyces pannorum* sensu lato and bat white nose syndrome pathogen *Geomyces (Pseudogymnoascus) destructans*. *Genome Announcements,* **1**, (6): e01045-13.
Childs-Sanford, S., Garner, M., Raymond, J., Didier, E. & Kollias, G. 2006. Disseminated microsporidiosis due to *Encephalitozoon hellem* in an Egyptian fruit bat (*Rousettus aegyptiacus*). *Journal of Comparative Pathology,* **134**, (4): 370-373.
Chroňáková, A., Horák, A., Elhottová, D. & Krištůfek, V. 2009. Diverse archaeal community of a bat guano pile in Domica Cave (Slovak Karst, Slovakia). *Folia Microbiologica,* **54**, (5): 436-446.

Chung, C.-U., Han, S.-H., Cha, J.-Y., Kim, S.-C., Kim, J.-J., Jeong, J.-C. & Lim, C.-W. 2015. The diet composition of the serotine bat, *Eptesicus serotinus* revealed by faecal analysis. *Korean Journal of Environmental Ecology,* **29**, (3): 368-373.
Clare, E. L., Barber, B. R., Sweeney, B. W., Hebert, P. D. N. & Fenton, M. B. 2011. Eating local: influences of habitat on the diet of little brown bats (*Myotis lucifugus*). *Molecular Ecology,* **20**, (8): 1772-1780.
Clare, E. L., Fraser, E. E., Braid, H. E., Fenton, M. B. & Hebert, P. D. N. 2009. Species on the menu of a generalist predator, the eastern red bat (*Lasiurus borealis*): using a molecular approach to detect arthropod prey. *Molecular Ecology,* **18**, (11): 2532-2542.
Clare, E. L., Lim, B. K., Engstrom, M. D., Eger, J. L. & Hebert, P. D. 2007. DNA barcoding of Neotropical bats: species identification and discovery within Guyana. *Molecular Ecology Notes,* **7**, (2): 184-190.
Clare, E. L., Symondson, W. O. & Fenton, M. B. 2014. An inordinate fondness for beetles? Variation in seasonal dietary preferences of night‐roosting big brown bats (*Eptesicus fuscus*). *Molecular Ecology,* **23**, (15): 3633-3647.
Clemente, J. C., Ursell, L. K., Parfrey, L. W. & Knight, R. 2012. The impact of the gut microbiota on human health: an integrative view. *Cell,* **148**, (6): 1258-1270.
Cody, R. 1989. Distribution of chitinase and chitobiase in Bacillus. *Current Microbiology,* **19**, (4): 201-205.
Coles, R. B., Guppy, A., Anderson, M. E. & Schlegel, P. 1989. Frequency sensitivity and directional hearing in the gleaning bat, *Plecotus auritus* (Linnaeus, 1758). *Journal of Comparative Physiology A,* **165**, (2): 269-280.
Concannon, R., Wynn-Owen, K., Simpson, V. & Birtles, R. 2005. Molecular characterization of haemoparasites infecting bats (Microchiroptera) in Cornwall, UK. *Parasitology,* **131**, (04): 489-496.
Conservation of Habitats and Species Regulations 2010. *London: The Stationery Office.*
Cooper, A. 1992. Removal of colourings, inhibitors of PCR, and the carrier effect of PCR contamination from ancient DNA samples. *Ancient DNA Newsletter,* **1**, (1): 31-32.
Corcoran, A. J. & Conner, W. E. 2012. Sonar jamming in the field: effectiveness and behavior of a unique prey defense. *The Journal of Experimental Biology,* **215**, (24): 4278-4287.
Corcoran, A. J. & Hristov, N. I. 2014. Convergent evolution of anti-bat sounds. *Journal of Comparative Physiology A,* **200**, (9): 811-821.
Cornish-Bowden, A. 1985. IUPAC-IUB symbols for nucleotide nomenclature. *Nucleic Acids Research,* **13**: 3021-3030.
Costanza, R., D'arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'neill, R. V. & Paruelo, J. 1997. The value of the world's ecosystem services and natural capital. *Nature,* **387**, (6630): 253-260.
Cui, J., Tachedjian, M., Wang, L., Tachedjian, G., Wang, L.-F. & Zhang, S. 2012. Discovery of retroviral homologs in bats: implications for the origin of mammalian gammaretroviruses. *Journal of Virology,* **86**, (8): 4288-4293.
Dale, C. & Moran, N. A. 2006. Molecular interactions between bacterial symbionts and their hosts. *Cell,* **126**, (3): 453-465.
Danko, Š., Krištín, A. & Krištofík, J. 2010. *Myotis alcathoe* in eastern Slovakia: Occurrence, diet, ectoparasites and notes on its identification in the field. *Vespertilio,* **13**: 77-91.
Davidson‐Watts, I. & Jones, G. 2006. Differences in foraging behaviour between *Pipistrellus pipistrellus* (Schreber, 1774) and *Pipistrellus pygmaeus* (Leach, 1825). *Journal of Zoology,* **268**, (1): 55-62.
De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E. & Taberlet, P. 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources,* **14**, (2): 306-323.
De Mandal, S., Panda, A. K., Bisht, S. S. & Kumar, N. S. 2015. First report of bacterial community from a bat guano using Illumina next-generation sequencing. *Genomics Data,* **4**: 99-101.

Deagle, B., Tollit, D., Jarman, S., Hindell, M., Trites, A. & Gales, N. 2005. Molecular scatology as a tool to study diet: analysis of prey DNA in scats from captive Steller sea lions. *Molecular Ecology,* **14**, (6): 1831-1842.
Deagle, B. E., Eveson, J. P. & Jarman, S. N. 2006. Quantification of damage in DNA recovered from highly degraded samples–a case study on DNA in faeces. *Frontiers in Zoology,* **3**, (1): 11.
Degnan, P. H., Yu, Y., Sisneros, N., Wing, R. A. & Moran, N. A. 2009. *Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. *Proceedings of the National Academy of Sciences,* **106**, (22): 9063-9068.
Dix, N. J. 2012. Fungal ecology*, Springer Science & Business Media.*
Don, R., Cox, P., Wainwright, B., Baker, K. & Mattick, J. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research,* **19**, (14): 4008.
Doyle, J. & Doyle, J. 1987. A rapid procedure for DNA purification from small quantities of fresh leaf tissue. *Phytochemical Bulletin,* **19**: 11-15.
Dray, S. & Dufour, A.-B. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software,* **22**, (4): 1-20.
Drexler, J. F., Geipel, A., König, A., Corman, V. M., Van Riel, D., Leijten, L. M., Bremer, C. M., Rasche, A., Cottontail, V. M. & Maganga, G. D. 2013. Bats carry pathogenic Hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. *Proceedings of the National Academy of Sciences,* **110**, (40): 16151-16156.
Eggert, L. S., Maldonado, J. E. & Fleischer, R. C. 2005. Nucleic acid isolation from ecological samples—animal scat and other associated materials. *Methods in Enzymology,* **395**: 73-82.
Emerson, G. L., Nordhausen, R., Garner, M. M., Huckabee, J. R., Johnson, S., Wohrle, R. D., Davidson, W. B., Wilkins, K., Li, Y. & Doty, J. B. 2013. Novel poxvirus in big brown bats, northwestern United States. *Emerging Infectious Diseases,* **19**: 1002-1004.
Emerson, J. K. & Roark, A. M. 2007. Composition of guano produced by frugivorous, sanguivorous, and insectivorous bats. *Acta Chiropterologica,* **9**, (1): 261-267.
Emmons, C. W., Klite, P., Baer, G. & Hill Jr, W. 1966. Isolation of *Histoplasma capsulatum* from bats in the United States. *American Journal of Epidemiology,* **84**, (1): 103-9.
Falk, J. J., Ter Hofstede, H. M., Jones, P. L., Dixon, M. M., Faure, P. A., Kalko, E. K. & Page, R. A. 2015. Sensory-based niche partitioning in a multiple predator–multiple prey community. *Proceedings of the Royal Society of London B: Biological Sciences,* **282**, (1808): 20150520.
Fang, J., Wang, X., Mu, S., Zhang, S. & Dong, D. 2015. BGD: a database of bat genomes. *PLoS ONE,* **10**, (6): e0131296.
Feldman, R., Whitaker, J. O. & Yom-Tov, Y. 2000. Dietary composition and habitat use in a desert insectivorous bat community in Israel. *Acta Chiropterologica,* **2**, (1): 15-22.
Fennestad, K. & Borg-Petersen, C. 1972. Leptospirosis in Danish wild mammals. *Journal of Wildlife Diseases,* **8**, (4): 343-351.
Fenton, M., Audet, D., Orbrist, M. & Rydell, J. 1995. Signal strength, timing, and self-deafening: the evolution of echolocation in bats. *Paleobiology,* **21**, (2): 229-242.
Fenton, M. B. 1997. Science and the conservation of bats. *Journal of Mammalogy,* **78**, (1): 1-14.
Fenton, M. B. & Bogdanowicz, W. 2002. Relationships between external morphology and foraging behaviour: bats in the genus Myotis. *Canadian Journal of Zoology,* **80**, (6): 1004-1013.
Fenton, M. B., Faure, P. A. & Ratcliffe, J. M. 2012. Evolution of high duty cycle echolocation in bats. *The Journal of Experimental Biology,* **215**, (17): 2935-2944.
Finke, M. D. 2002. Complete nutrient composition of commercially raised invertebrates used as food for insectivores. *Zoo Biology,* **21**, (3): 269-285.

Flanders, J. & Jones, G. 2009. Roost use, ranging behavior, and diet of greater horseshoe bats (*Rhinolophus ferrumequinum*) using a transitional roost. *Journal of Mammalogy,* **90**, (4): 888-896.
Flaquer, C., Puig-Montserrat, X., Goiti, U., Vidal, F., Curcó, A. & Russo, D. 2009. Habitat selection in Nathusius' pipistrelle (*Pipistrellus nathusii*): the importance of wetlands. *Acta Chiropterologica,* **11**, (1): 149-155.
Flavin, D. A., Biggane, S. S., Shiel, C. B., Smiddy, P. & Fairley, J. S. 2001. Analysis of the diet of Daubenton's bat *Myotis daubentonii* in Ireland. *Acta Theriologica,* **46**, (1): 43-52.
Flint, H. J., Scott, K. P., Duncan, S. H., Louis, P. & Forano, E. 2012. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes,* **3**, (4): 289-306.
Florens, F. V. 2015. Flying foxes face cull despite evidence. *Science,* **350**, (6266): 1325-1325.
França, D., Pereira, S., Maas, A., Martins, M., Bolzan, D., Lima, I., Dias, D. & Peracchi, A. 2013. Ectoparasitic flies (Diptera, Streblidae) of bats (Chiroptera, Phyllostomidae) in an Atlantic Forest area, southeastern Brazil. *Brazilian Journal of Biology,* **73**, (4): 847-854.
Fredricks, D. N., Smith, C. & Meier, A. 2005. Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. *Journal of Clinical Microbiology,* **43**, (10): 5122-5128.
Freeman, P. W. 1979. Specialized insectivory: beetle-eating and moth-eating molossid bats. *Journal of Mammalogy,* **60**, (3): 467-479.
Freeman, P. W. 1981. Correspondence of food habits and morphology in insectivorous bats. *Journal of Mammalogy,* **62**, (1): 166-173.
Freeman, P. W. 1998. Form, function, and evolution in skulls and teeth of bats*, Bat Biology and Conservation.* 140-156.
Freeman, P. W. & Lemen, C. A. 2007. Using scissors to quantify hardness of insects: do bats select for size or hardness? *Journal of Zoology,* **271**, (4): 469-476.
Gajdosik, M. & Gaisler, J. 2004. Diet of two Eptesicus bat species in Moravia (Czech Republic). *Folia Zoologica,* **53**, (1): 7-16.
Galante, E. & Marcos-Garcia, M. A. 2004. Decomposer insects. *Encyclopedia of Entomology.* Springer.
Garcia-Solache, M. A. & Casadevall, A. 2010. Global warming will bring new fungal diseases for mammals. *MBio,* **1**, (1): e00061-10.
Gaston, K. J. & O'neill, M. A. 2004. Automated species identification: why not? *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* **359**, (1444): 655-667.
Ge, X., Li, Y., Yang, X., Zhang, H., Zhou, P., Zhang, Y. & Shi, Z. 2012. Metagenomic analysis of viruses from the bat fecal samples reveals many novel viruses in insectivorous bats in China. *Journal of Virology*: JVI. 06671-11.
Geiser, F. & Ruf, T. 1995. Hibernation versus daily torpor in mammals and birds: physiological variables and classification of torpor patterns. *Physiological Zoology*: 935-966.
Geisler, H. & Dietz, M. 1999. Zur nahrungsökologie einer wochenstubenkolonie der fransenfledermaus (*Myotis nattereri* Kuhl, 1818) in Mittelhessen. *Nyctalus,* **7**, (1): 87-101.
Gelhaus, M. & Zahn, A. 2010. Roosting ecology, phenology and foraging habitat of a nursery colony of *Pipistrellus nathusii* in the southwestern part of its reproduction range. *Vespertilio,* **13**, (14): 93-102.
Gerber, E., Haffner, M. & Ziswiler, V. 1994. Comparative food analysis of the serotine bat *Eptesinus serotinus* (Schreber, 1774) (Mammalia, Chiroptera) in different regions of Switzerland. *Revue Suisse de Zoologie,* **101**, (4): 858.
Ghazali, M. & Dzeverin, I. 2013. Correlations between hardness of food and craniodental traits in nine Myotis species (Chiroptera, Vespertilionidae). *Vestnik Zoologii,* **47**, (1): 73-8.
Gibson, K. E., Rikihisa, Y., Zhang, C. & Martin, C. 2005. *Neorickettsia risticii* is vertically transmitted in the trematode *Acanthatrium oregonense* and horizontally transmitted to bats. *Environmental Microbiology,* **7**, (2): 203-212.

Gill, P. G., Purnell, M. A., Crumpton, N., Brown, K. R., Gostling, N. J., Stampanoni, M. & Rayfield, E. J. 2014. Dietary specializations and diversity in feeding ecology of the earliest stem mammals. *Nature,* **512**, (7514): 303-305.
Ginolhac, A., Rasmussen, M., Gilbert, M. T. P., Willerslev, E. & Orlando, L. 2011. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics,* **27**, (15): 2153-2155.
Giovino, A., Marino, P., Domina, G., Scialabba, A., Schicchi, R., Diliberto, G., Rizza, C. & Scibetta, S. 2014. Evaluation of the DNA barcoding approach to develop a reference data-set for the threatened flora of Sicily. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*: 1-10.
Gloor, S., Stutz, H.-P. B. & Ziswiler, V. 1989. Nutritional habits of the noctule bat *Nyctalus noctula* (Schreber, 1714) in Switzerland. *Myotis,* **32**: 231-242.
Golenberg, E. M., Bickel, A. & Weihs, P. 1996. Effect of highly fragmented DNA on PCR. *Nucleic Acids Research,* **24**, (24): 5026-5033.
Gomez-Alvarez, V., Teal, T. K. & Schmidt, T. M. 2009. Systematic artifacts in metagenomes from complex microbial communities. *The ISME Journal,* **3**, (11): 1314-1317.
Goodfellow, M. & Williams, S. 1983. Ecology of Actinomycetes. *Annual Reviews in Microbiology,* **37**, (1): 189-216.
Gordon, A. & Hannon, G. 2010. Fastx-toolkit. FASTQ/A short-reads pre-processing tools. *Unpublished Available online at:* http://hannonlab/*. cshl. edu/fastx_toolkit*.
Greenberg, B. 1998. Reproductive status of some overwintering domestic flies (Diptera: Muscidae and Calliphoridae) with forensic implications. *Annals of the Entomological Society of America,* **91**, (6): 818-820.
Greenwood, J. J., Gregory, R. D., Harris, S., Morris, P. A. & Yalden, D. W. 1996. Relations between abundance, body size and species number in British birds and mammals. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* **351**, (1337): 265-278.
Gregory, T. & Shorthouse, D. 2003. Genome sizes of spiders. *Journal of Heredity,* **94**, (4): 285-290.
Griffiths, D. 1975. Prey availability and the food of predators. *Ecology,* **56**: 1209-1214.
Grinnell, J. 1917. The niche-relationships of the California Thrasher. *The Auk,* **34**, (4): 427-433.
Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W. & Hebert, P. D. 2006. DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America,* **103**, (4): 968-971.
Hajkova, P. & Pikula, J. 2007. Veterinary treatment of evening bats (Vespertilionidae) in the Czech Republic. *Veterinary Record,* **161**, (4): 139.
Hallam, T. G. & Mccracken, G. F. 2011. Management of the panzootic white‐nose syndrome through culling of bats. *Conservation Biology,* **25**, (1): 189-194.
Han, H.-J., Wen, H.-L., Zhou, C.-M., Chen, F.-F., Luo, L.-M., Liu, J.-W. & Yu, X.-J. 2015. Bats as reservoirs of severe emerging infectious diseases. *Virus Research,* **205**: 1-6.
Handelsman, J. 2004. Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews,* **68**, (4): 669-685.
Hanson, A. W. 1970. Isolation of spirochaetes from primates and other mammalian species. *The British Journal of Venereal Diseases,* **46**, (4): 303-306.
Hardin, G. 1960. The competitive exclusion principle. *Science,* **131**, (3409): 1292-1297.
Hazeleger, W., Jacobs-Reitsma, W., Van Bergen, M., Lina, P., Van Den Bosch, T. & Beumer, R. Bats may be carriers of Campylobacter. 12th European Bat Research Symposium, Vilnius, Lithuania, August 22-26, 2011, 2011. 66-67.
Hebert, P. D. & Gregory, T. R. 2005. The promise of DNA barcoding for taxonomy. *Systematic Biology,* **54**, (5): 852-859.
Hebert, P. D., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America,* **101**, (41): 14812-14817.

Hebert, P. D., Ratnasingham, S. & De Waard, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences,* **270**, (Suppl 1): S96-S99.

Heck Jr, K. L., Van Belle, G. & Simberloff, D. 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. *Ecology*: 1459-1461.

Heinicke, W. & Krauß, A. 1978. Zum beutespektrum des braunen langohrs, *Plecotus auritus*. *Nyctalus (NF),* **1**: 49-52.

Hert, D. G., Fredlake, C. P. & Barron, A. E. 2008. Advantages and limitations of next‐generation sequencing technologies: A comparison of electrophoresis and non‐electrophoresis methods. *Electrophoresis,* **29**, (23): 4618-4626.

Hill, J. E. 1974. A new family, genus and species of bat (Mammalia: Chiroptera) from Thailand. *Bulletin of the British Museum (Natural History), Zoology,* **27**: 301-336.

Ho, Y.-Y., Fang, Y.-P., Chou, C.-H., Cheng, H.-C. & Chang, H.-W. 2013. High duty cycle to low duty cycle: echolocation behaviour of the hipposiderid bat *Coelops frithii*. *PLoS ONE,* **8**, (5): e62938.

Hoare, L. 1991. The diet of *Pipistrellus pipistrellus* during the pre‐hibernal period. *Journal of Zoology,* **225**, (4): 665-670.

Hofreiter, M., Paijmans, J. L., Goodchild, H., Speller, C. F., Barlow, A., Fortes, G. G., Thomas, J. A., Ludwig, A. & Collins, M. J. 2015. The future of ancient DNA: Technical advances and conceptual shifts. *BioEssays,* **37**, (3): 284-293.

Holderied, M. & Von Helversen, O. 2003. Echolocation range and wingbeat period match in aerial-hawking bats. *Proceedings of the Royal Society of London B: Biological Sciences,* **270**, (1530): 2293-2299.

Hoofer, S. R., Reeder, S. A., Hansen, E. W. & Van Den Bussche, R. A. 2003. Molecular phylogenetics and taxonomic review of noctilionoid and vespertilionoid bats (Chiroptera: Yangochiroptera). *Journal of Mammalogy,* **84**, (3): 809-821.

Hope, P. R., Bohmann, K., Gilbert, M. T. P., Zepeda-Mendoza, M. L., Razgour, O. & Jones, G. 2014. Second generation sequencing and morphological faecal analysis reveal unexpected foraging behaviour by *Myotis nattereri* (Chiroptera, Vespertilionidae) in winter. *Frontiers in Zoology,* **11**, (1): 39.

Hopkins, J. B. & Ferguson, J. M. 2012. Estimating the diets of animals using stable isotopes and a comprehensive Bayesian mixing model. *PLoS ONE,* **7**, (1): e28478.

Houbraken, J., Frisvad, J. C. & Samson, R. A. 2011. Fleming's penicillin producing strain is not *Penicillium chrysogenum* but *P. rubens*. *IMA Fungus,* **2**, (1): 87-95.

Hulva, P., Horáček, I., Strelkov, P. P. & Benda, P. 2004. Molecular architecture of *Pipistrellus pipistrellus/Pipistrellus pygmaeus* complex (Chiroptera: Vespertilionidae): further cryptic species and Mediterranean origin of the divergence. *Molecular Phylogenetics and Evolution,* **32**, (3): 1023-1035.

Hundt, L. 2012. Bat surveys: good practice guidelines*, Bat Conservation Trust.* 11-15.

Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. 2007. MEGAN analysis of metagenomic data. *Genome Research,* **17**, (3): 377-386.

Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N. & Schuster, S. C. 2011. Integrative analysis of environmental sequences using MEGAN4. *Genome Research,* **21**, (9): 1552-1560.

Hutson, A. M., Spitzenberger, F., Aulagnier, S., Coroiu, I., Karataş, A., Juste, J., Paunovic, M., Palmeirim, J. & Benda, P. 2008. *Myotis myotis* [Online]. Available: http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T14133A4403709.en [Accessed 30 November 2015].

I.C.Z.N. 2003. Opinion 2028 (Case 3073). *Vespertilio pipistrellus,* Schreber, 1774 and *V. pygmaeus,* Leach, 1825 (currently *Pipistrellus pipistrellus* and *P. pygmaeus*; Mammalia, Chiroptera): neotypes designated. *Bulletin of Zoological Nomenclature,* **60**, (56): 182-186.

I.U.C.N. 2013. *IUCN red list of threatened species* [Online]. Available: http://www.iucnredlist.org/ [Accessed 30 November 2015].

Idaghdour, Y., Broderick, D. & Korrida, A. 2003. Faeces as a source of DNA for molecular studies in a threatened population of great bustards. *Conservation Genetics,* **4**, (6): 789-792.
Illumina. 2013a. *Nextera DNA sample preparation kits* [Online]. Available: http://www.illumina.com/products/nextera_dna_sample_prep_kit.ilmn 2013].
Illumina. 2013b. *TruSeq DNA sample preparation kits* [Online]. Available: http://www.illumina.com/products/truseq_dna_sample_prep_kits.ilmn [Accessed 17th June 2013].
Islas, S., Becerra, A., Luisi, P. L. & Lazcano, A. 2004. Comparative genomics and the gene complement of a minimal cell. *Origins of Life and Evolution of the Biosphere,* **34**, (1-2): 243-256.
Jan, C. M., Frith, K., Glover, A. M., Butlin, R. K., Scott, C. D., Greenaway, F., Ruedi, M., Frantz, A. C., Dawson, D. A. & Altringham, J. D. 2010. *Myotis alcathoe* confirmed in the UK from mitochondrial and microsatellite DNA. *Acta Chiropterologica,* **12**, (2): 471-483.
Jedlicka, J. A., Sharma, A. M. & Almeida, R. P. 2013. Molecular tools reveal diets of insectivorous birds from predator fecal matter. *Conservation Genetics Resources,* **5**, (3): 879-885.
Jeuniaux, C. 1961. Chitinase: an addition to the list of hydrolases in the digestive tract of vertebrates. *Nature,* **192**, (4798): 135-136.
Jin, L., Feng, J., Sun, K., Liu, Y., Wu, L., Li, Z. & Zhang, X. 2005. Foraging strategies in the greater horseshoe bat (*Rhinolophus ferrumequinum*) on Lepidoptera in summer. *Chinese Science Bulletin,* **50**, (14): 1477-1482.
Jolliffe, I. 1986. Principal component analysis*, New York: Springer-Verlag.*
Jonason, D., Franzén, M. & Ranius, T. 2014. Surveying moths using light traps: effects of weather and time of year. *PLoS ONE,* **9**, (3): e92453.
Jones, G. 1990. Prey selection by the greater horseshoe bat (*Rhinolophus ferrumequinum*): optimal foraging by echolocation? *The Journal of Animal Ecology*: 587-602.
Jones, G. Flight morphology, flight performance and echolocation in British bats. Proceedings of the 1st European bat detector workshop. Netherlands Bat Research Foundation, Amsterdam, 1993. 59-78.
Jones, G. 1995. Flight performance, echolocation and foraging behaviour in noctule bats *Nyctalus noctula*. *Journal of Zoology,* **237**, (2): 303-312.
Jones, G. & Barratt, E. 1999. *Vespertilio pipistrellus* Schreber, 1774 and *V. pygmaeus* Leach, 1825 (currently *Pipistrellus pipistrellus* and *P. pygmaeus*; Mammalia, Chiroptera): proposed designation of neotypes. *Bulletin of Zoological Nomenclature,* **56**: 182-186.
Jones, G., Duverge, P. L. & Ransome, R. D. Conservation biology of an endangered species: field studies of greater horseshoe bats. Symposia of the Zoological Society of London, 1995. London: The Society, 1960-1999, 309-324.
Jones, G., Jacobs, D. S., Kunz, T. H., Willig, M. R. & Racey, P. A. 2009. Carpe noctem: the importance of bats as bioindicators. *Endangered Species Research,* **8**, (1-2): 93-115.
Jones, G. & Rayner, J. 1988. Flight performance, foraging tactics and echolocation in free‐living Daubenton's bats *Myotis daubentoni* (Chiroptera: Vespertilionidae). *Journal of Zoology,* **215**, (1): 113-132.
Jones, G. & Rayner, J. M. 1989. Foraging behavior and echolocation of wild horseshoe bats *Rhinolophus ferrumequinum* and *R. hipposideros* (Chiroptera, Rhinolophidae). *Behavioral Ecology and Sociobiology*, **25**, (3): 183-191.
Jones, G. & Rydell, J. 1994. Foraging strategy and predation risk as factors influencing emergence time in echolocating bats. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **346**, (1318): 445-455.
Jones, G. & Teeling, E. C. 2006. The evolution of echolocation in bats. *Trends in Ecology & Evolution,* **21**, (3): 149-156.
Jones, K. E. & Maclarnon, A. 2001. Bat life histories: testing models of mammalian life-history evolution. *Evolutionary Ecology Research,* **3**, (4): 487-505.

Jones, K. E., Purvis, A. & Gittleman, J. L. 2003. Biological correlates of extinction risk in bats. *The American Naturalist,* **161**, (4): 601-614.
Jones, K. E., Purvis, A., Maclarnon, A., Bininda‐Emonds, O. R. & Simmons, N. B. 2002. A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biological Reviews,* **77**, (2): 223-259.
Kalko, E. K. & Schnitzler, H.-U. 1989. The echolocation and hunting behavior of Daubenton's bat, *Myotis daubentoni*. *Behavioral Ecology and Sociobiology,* **24**, (4): 225-238.
Kalko, E. K. & Schnitzler, H.-U. 1993. Plasticity in echolocation signals of European pipistrelle bats in search flight: implications for habitat use and prey detection. *Behavioral Ecology and Sociobiology,* **33**, (6): 415-428.
Kaňuch, P., Janečková, K. & Krištín, A. 2005a. Winter diet of the noctule bat *Nyctalus noctula*. *Folia Zoologica,* **54**, (1-2): 53-60.
Kaňuch, P., Krištín, A. & Krištofík, J. 2005b. Phenology, diet, and ectoparasites of Leisler's bat (*Nyctalus leisleri*) in the Western Carpathians (Slovakia). *Acta Chiropterologica,* **7**, (2): 249-257.
Kavanagh, K. & Reeves, E. P. 2004. Exploiting the potential of insects for in vivo pathogenicity testing of microbial pathogens. *FEMS Microbiology Reviews,* **28**, (1): 101-112.
Kervyn, T. & Libois, R. 2008. The diet of the serotine bat: A comparison between rural and urban environments. *Belgian Journal of Zoology,* **138**, (1).
Khan, G., Kangro, H., Coates, P. & Heath, R. 1991. Inhibitory effects of urine on the polymerase chain reaction for cytomegalovirus DNA. *Journal of Clinical Pathology,* **44**, (5): 360-365.
King, A. M. 2011. Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses*, Elsevier.* 9.
Kistler, L., Smith, O., Ware, R., Momber, G., Bates, R., Garwood, P., Fitch, S., Pallen, M., Gaffney, V. & Allaby, R. 2015. Thermal age, cytosine deamination and the veracity of 8,000 year old wheat DNA from sediments. *bioRxiv*: 032060.
Klingen, I., Eilenberg, J. & Meadow, R. 2002. Effects of farming system, field margins and bait insect on the occurrence of insect pathogenic fungi in soils. *Agriculture, Ecosystems & Environment,* **91**, (1): 191-198.
Knapp, M. & Hofreiter, M. 2010. Next generation sequencing of ancient DNA: requirements, strategies and perspectives. *Genes,* **1**, (2): 227-243.
Knierim, E., Lucke, B., Schwarz, J. M., Schuelke, M. & Seelow, D. 2011. Systematic comparison of three methods for fragmentation of long-range PCR products for next generation sequencing. *PLoS ONE,* **6**, (11): e28240.
Kobayashi, N., Tamura, K. & Aotsuka, T. 1999. PCR error and molecular population genetics. *Biochemical Genetics,* **37**, (9): 317-321.
Köchl, S., Niederstätter, H. & Parson, W. 2005. DNA extraction and quantitation of forensic samples using the phenol-chloroform method and real-time PCR. *Forensic DNA Typing Protocols.* Springer.
Kohl, C., Vidovszky, M. Z., Mühldorfer, K., Dabrowski, P. W., Radonić, A., Nitsche, A., Wibbelt, G., Kurth, A. & Harrach, B. 2011. Genome analysis of bat adenovirus 2: indications of interspecies transmission. *Journal of Virology*: JVI. 05974-11.
Korbie, D. J. & Mattick, J. S. 2008. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols,* **3**, (9): 1452-1456.
Kosoy, M., Bai, Y., Lynch, T., Kuzmin, I. V., Niezgoda, M., Franka, R., Agwanda, B., Breiman, R. F. & Rupprecht, C. E. 2010. Bartonella spp. in bats, Kenya. *Emerging Infectious Diseases,* **16**, (12): 1875-81.
Kozarewa, I., Ning, Z., Quail, M. A., Sanders, M. J., Berriman, M. & Turner, D. J. 2009. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+ C)-biased genomes. *Nature Methods,* **6**, (4): 291-295.
Krauel, J. J., Westbrook, J. K. & Mccracken, G. F. 2015. Weather‐driven dynamics in a dual‐migrant system: moths and bats. *Journal of Animal Ecology,* **84**, (3): 604-614.
Krauss, A. 1978. Materials on foodstuff biology of long-eared bat *Plecotus auritus*, Mammalia, Chiroptera. *Zoologische Abhandlungen (Dresden),* **34**, (22): 325-338.

Krüger, F., Clare, E., Greif, S., Siemers, B., Symondson, W. & Sommer, R. 2013a. An integrative approach to detect subtle trophic niche differentiation in the sympatric trawling bat species *Myotis dasycneme* and *Myotis daubentonii*. *Molecular Ecology,* **23**, (15): 3657-3671.

Krüger, F., Clare, E., Greif, S., Siemers, B., Symondson, W. O. C. & Sommer, R. 2014. An integrative approach to detect subtle trophic niche differentiation in the sympatric trawling bat species *Myotis dasycneme* and *Myotis daubentonii*. *Molecular Ecology,* **23**, (15): 3657-3671.

Krüger, F., Clare, E. L., Symondson, W. O., Keišs, O. & Pētersons, G. 2013b. Diet of the insectivorous bat *Pipistrellus nathusii* during autumn migration and summer residence. *Molecular Ecology,* **23**, (15): 3672-3683.

Krüger, F., Harms, I., Fichtner, A., Wolz, I. & Sommer, R. S. 2012. High trophic similarity in the sympatric North European trawling bat species *Myotis daubentonii* and *Myotis dasycneme*. *Acta Chiropterologica,* **14**, (2): 347-356.

Kuddus, M. & Ahmad, I. 2013. Isolation of novel chitinolytic bacteria and production optimization of extracellular chitinase. *Journal of Genetic Engineering and Biotechnology,* **11**, (1): 39-46.

Kunz, T., Whitaker Jr, J. & Wadanoli, M. 1995. Dietary energetics of the insectivorous Mexican free-tailed bat (*Tadarida brasiliensis*) during pregnancy and lactation. *Oecologia,* **101**, (4): 407-415.

Kunz, T. H., Braun De Torrez, E., Bauer, D., Lobova, T. & Fleming, T. H. 2011. Ecosystem services provided by bats. *Annals of the New York Academy of Sciences,* **1223**, (1): 1-38.

Lagerveld, S., Poerink, B. J., Haselager, R. & Verdaat, H. 2014. Bats in Dutch offshore wind farms in autumn 2012. *Lutra,* **57**, (2): 61-69.

Lahr, D. & Katz, L. A. 2009. Reducing the impact of PCR-mediated recombination in molecular evolution and environmental studies using a new-generation high-fidelity DNA polymerase. *Biotechniques,* **47**, (4): 857-866.

Lantz, P.-G., Matsson, M., Wadström, T. & Rådström, P. 1997. Removal of PCR inhibitors from human faecal samples through the use of an aqueous two-phase system for sample preparation prior to PCR. *Journal of Microbiological Methods,* **28**, (3): 159-167.

Leopardi, S., Blake, D. & Puechmaille, S. J. 2015. White-nose syndrome fungus introduced from Europe to North America. *Current Biology,* **25**, (6): R217-R219.

Lepp, D., Roxas, B., Parreira, V. R., Marri, P. R., Rosey, E. L., Gong, J., Songer, J. G., Vedantam, G. & Prescott, J. F. 2010. Identification of novel pathogenicity loci in *Clostridium perfringens* strains that cause avian necrotic enteritis. *PLoS ONE,* **5**, (5): e10795.

Levin, E., Yom-Tov, Y., Hefetz, A. & Kronfeld-Schor, N. 2013. Changes in diet, body mass and fatty acid composition during pre-hibernation in a subtropical bat in relation to NPY and AgRP expression. *Journal of Comparative Physiology B,* **183**, (1): 157-166.

Levins, R. 1968. Evolution in changing environments: some theoretical explorations.(MPB-2)*, Princeton University Press.* 2.

Li, G., Wang, J., Rossiter, S. J., Jones, G., Cotton, J. A. & Zhang, S. 2008. The hearing gene Prestin reunites echolocating bats. *Proceedings of the National Academy of Sciences,* **105**, (37): 13959-13964.

Li, L., Victoria, J. G., Wang, C., Jones, M., Fellers, G. M., Kunz, T. H. & Delwart, E. 2010. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *Journal of Virology,* **84**, (14): 6955-6965.

Li, Y., Liu, Y., Zhou, Z., Huang, H., Ren, Y., Zhang, Y., Li, G., Zhou, Z. & Wang, L. 2011. Complete genome sequence of *Aeromonas veronii* strain B565. *Journal of Bacteriology,* **193**, (13): 3389-3390.

Lin, J.-W., Hsu, Y.-M., Chomel, B. B., Lin, L.-K., Pei, J.-C., Wu, S.-H. & Chang, C.-C. 2012. Identification of novel Bartonella spp. in bats and evidence of Asian gray shrew as a new potential reservoir of Bartonella. *Veterinary Microbiology,* **156**, (1): 119-126.

Lino, A., Fonseca, C., Goiti, U. & Pereira, M. J. R. 2014. Prey selection by *Rhinolophus hipposideros* (Chiroptera, Rhinolophidae) in a modified forest in Southwest Europe. *Acta Chiropterologica,* **16**, (1): 75-83.

López-Hoffman, L., Wiederholt, R., Sansone, C., Bagstad, K. J., Cryan, P., Diffendorfer, J. E., Goldstein, J., Lasharr, K., Loomis, J. & Mccracken, G. 2014. Market forces and technological substitutes cause fluctuations in the value of bat pest-control services for cotton. *PLoS ONE,* **9**, (2): e87912.

Lorch, J. M., Lindner, D. L., Gargas, A., Muller, L. K., Minnis, A. M. & Blehert, D. S. 2012. A culture-based survey of fungi in soil from bat hibernacula in the eastern United States and its implications for detection of *Geomyces destructans*, the causal agent of bat white-nose syndrome. *Mycologia*: 12-207.

Lorch, J. M., Meteyer, C. U., Behr, M. J., Boyles, J. G., Cryan, P. M., Hicks, A. C., Ballmann, A. E., Coleman, J. T., Redell, D. N. & Reeder, D. M. 2011. Experimental infection of bats with *Geomyces destructans* causes white-nose syndrome. *Nature,* **480**, (7377): 376-378.

Lourenço, E. C., Patrício, P. M. P., Pinheiro, M. D. C., Dias, R. M. & Famadas, K. M. 2014. Streblidae (Diptera) on bats (Chiroptera) in an area of Atlantic Forest, state of Rio de Janeiro. *Revista Brasileira de Parasitologia Veterinária,* **23**, (2): 164-170.

Lučan, R. K. 2009. Effect of colony size and reproductive period on the emergence behaviour of a maternity colony of Daubenton's Bat (*Myotis daubentonii*) occupying an artificial roost (Chiroptera: Vespertilionidae). *Lynx,* **40**: 71-81.

Lucan, R. K., Andreas, M., Benda, P., Bartonicka, T., Brezinová, T., Hoffmannová, A., Hulová, Š., Hulva, P., Neckarova, J. & Reiter, A. 2009. Alcathoe bat (*Myotis alcathoe*) in the Czech Republic: distributional status, roosting and feeding ecology. *Acta Chiropterologica,* **11**, (1): 61-69.

Lučan, R. K., Hanák, V. & Horáček, I. 2009. Long-term re-use of tree roosts by European forest bats. *Forest Ecology and Management,* **258**, (7): 1301-1306.

Luo, J., Koselj, K., Zsebők, S., Siemers, B. M. & Goerlitz, H. R. 2014. Global warming alters sound transmission: differential impact on the prey detection ability of echolocating bats. *Journal of The Royal Society Interface,* **11**, (91): e20130961.

Ma, J., Liang, B., Zhang, S. & Metzner, W. 2008. Dietary composition and echolocation call design of three sympatric insectivorous bat species from China. *Ecological Research,* **23**, (1): 113-119.

Mackenzie, G. & Oxford, G. 1995. Prey of the noctule bat (*Nyctalus noctula*) in East Yorkshire. *Journal of Zoology,* **236**, (2): 322-327.

Maine, J. J. & Boyles, J. G. 2015. Bats initiate vital agroecological interactions in corn. *Proceedings of the National Academy of Sciences,* **112**, (40): 12438-12443.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research,* **27**, (2 Part 1): 209-220.

Manwaring, J. 1939. Lepidoptera taken by bats. *Entomologist,* **72**: 190.

Mardis, E. R. 2008a. The impact of next-generation sequencing technology on genetics. *Trends in Genetics,* **24**, (3): 133-141.

Mardis, E. R. 2008b. Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics,* **9**: 387-402.

Markoulatos, P., Siafakas, N. & Moncany, M. 2002. Multiplex polymerase chain reaction: a practical approach. *Journal of Clinical Laboratory Analysis,* **16**, (1): 47-51.

Martens, E. C., Chiang, H. C. & Gordon, J. I. 2008. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host & Microbe,* **4**, (5): 447-457.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal,* **17**, (1): pp. 10-12.

Matthias, M. A., Díaz, M. M., Campos, K. J., Calderon, M., Willig, M. R., Pacheco, V., Gotuzzo, E., Gilman, R. H. & Vinetz, J. M. 2005. Diversity of bat-associated Leptospira in the Peruvian Amazon inferred by bayesian phylogenetic analysis of 16S ribosomal DNA sequences. *The American Journal of Tropical Medicine and Hygiene,* **73**, (5): 964-974.

Mayer, F., Dietz, C. & Kiefer, A. 2007. Molecular species identification boosts bat diversity. *Frontiers in Zoology,* **4**, (4): 5.

Mayer, F. & Von Helversen, O. 2001. Cryptic diversity in European bats. *Proceedings of the Royal Society of London B: Biological Sciences,* **268**, (1478): 1825-1832.

Mcaney, C. 1991. The analysis of bat droppings*, Mammal Society.*

Mcaney, C. & Fairley, J. 1989. Analysis of the diet of the lesser horseshoe bat *Rhinolophus hipposideros* in the West of Ireland. *Journal of Zoology,* **217**, (3): 491-498.
Meyer, M. & Kircher, M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols,* **2010**, (6): pdb. prot5448.
Mickleburgh, S. P., Hutson, A. M. & Racey, P. A. 2002. A review of the global conservation status of bats. *Oryx,* **36**, (1): 18-34.
Middleton, N., Froud, A. & French, K. 2014. Social calls of the bats of Britain and Ireland*, Exeter: Pelagic Publishing Ltd.* 1-179.
Mikheyev, A. S. & Tin, M. M. 2014. A first look at the Oxford Nanopore MinION sequencer. *Molecular Ecology Resources,* **14**, (6): 1097-1102.
Mikula, P. & Čmoková, A. 2012. Lepidopterans in the summer diet of *Eptesicus serotinus* in Central Bohemia. *Vespertilio,* **16**: 197–201.
Miller, J. D. 1992. Fungi as contaminants in indoor air. *Atmospheric Environment. Part A. General Topics,* **26**, (12): 2163-2172.
Miller, L. A. & Surlykke, A. 2001. How some insects detect and avoid being eaten by bats: tactics and countertactics of prey and predator. *BioScience,* **51**, (7): 570-581.
Monteiro, L., Bonnemaison, D., Vekris, A., Petry, K. G., Bonnet, J., Vidal, R., Cabrita, J. & Mégraud, F. 1997. Complex polysaccharides as PCR inhibitors in feces: *Helicobacter pylori* model. *Journal of Clinical Microbiology,* **35**, (4): 995-998.
Moran, N. A., Degnan, P. H., Santos, S. R., Dunbar, H. E. & Ochman, H. 2005a. The players in a mutualistic symbiosis: insects, bacteria, viruses, and virulence genes. *Proceedings of the National Academy of Sciences of the United States of America,* **102**, (47): 16919-16926.
Moran, N. A., Russell, J. A., Koga, R. & Fukatsu, T. 2005b. Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Applied and Environmental Microbiology,* **71**, (6): 3302-3310.
Moretti, T., Koons, B. & Budowle, B. 1998. Enhancement of PCR amplification yield and specificity using AmpliTaq Gold DNA polymerase. *Biotechniques,* **25**, (4): 716-722.
Moussian, B., Schwarz, H., Bartoszewski, S. & Nüsslein‐Volhard, C. 2005. Involvement of chitin in exoskeleton morphogenesis in *Drosophila melanogaster*. *Journal of Morphology,* **264**, (1): 117-130.
Mühldorfer, K. 2013. Bats and bacterial pathogens: a review. *Zoonoses and Public Health,* **60**, (1): 93-103.
Mühldorfer, K., Schwarz, S., Fickel, J., Wibbelt, G. & Speck, S. 2011a. Genetic diversity of Pasteurella species isolated from European vespertilionid bats. *Veterinary Microbiology,* **149**, (1): 163-171.
Mühldorfer, K., Speck, S., Kurth, A., Lesnik, R., Freuling, C., Müller, T., Kramer-Schadt, S. & Wibbelt, G. 2011b. Diseases and causes of death in European bats: dynamics in disease susceptibility and infection rates. *PLoS ONE,* **6**, (12): e29773.
Mulec, J., Covington, E. & Walochnik, J. 2013. Is bat guano a reservoir of *Geomyces destructans*? *Open Journal of Veterinary Medicine,* **3**, (2): 161-167.
Müller, F. 1878. Ueber die vortheile der mimicry bei schmetterlingen. *Zoologischer Anzeiger*, (1): 54-55.
N.C.B.I. Resource Coordinators 2013. Database resources of the national center for biotechnology information. *Nucleic Acids Research,* **41**, (Database issue): D8.
Nakagawa, S., Takaki, Y., Shimamura, S., Reysenbach, A.-L., Takai, K. & Horikoshi, K. 2007. Deep-sea vent ε-proteobacterial genomes provide insights into emergence of pathogens. *Proceedings of the National Academy of Sciences,* **104**, (29): 12146-12150.
Nelson, K., Weinel, C., Paulsen, I., Dodson, R., Hilbert, H., Martins Dos Santos, V., Fouts, D., Gill, S., Pop, M. & Holmes, M. 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environmental Microbiology,* **4**, (12): 799-808.
Neuweiler, G. 1990. Auditory adaptations for prey capture in echolocating bats. *Physiological Reviews,* **70**, (3): 615-641.

Nissen, H., Krueger, F., Fichtner, A. & Sommer, R. S. 2013. Local variability in the diet of Daubenton's bat (*Myotis daubentonii*) in a lake landscape of northern Germany. *Folia Zoologica,* **62**, (1): 36-41.
Norberg, U. 1995. How a long tail and changes in mass and wing shape affect the cost for flight in animals. *Functional Ecology,* **9**, (1): 48-54.
Nováková, A. 2009. Microscopic fungi isolated from the Domica Cave system (Slovak Karst National Park, Slovakia). A review. *International Journal of Speleology,* **38**, (1): 8.
Obrist, M. K., Boesch, R. & Flückiger, P. F. 2004. Variability in echolocation call design of 26 Swiss bat species: consequences, limits and options for automated field identification with a synergetic pattern recognition approach. *Mammalia,* **68**, (4): 307-322.
Oksanen, J. 2010. Cluster analysis: tutorial with R. *University of Oulu, Oulu*.
Oliver, K. M., Russell, J. A., Moran, N. A. & Hunter, M. S. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proceedings of the National Academy of Sciences,* **100**, (4): 1803-1807.
Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B. & Moltke, I. 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature,* **499**, (7456): 74-78.
Ostell, J. M., Wheelan, S. J. & Kans, J. A. 2004. The NCBI data model. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins,* **43**: 19.
Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. & Hofreiter, M. 2004. Genetic analyses from ancient DNA. *Annual Review of Genetics,* **38**: 645-679.
Palmer, S. A., Clapham, A. J., Rose, P., Freitas, F. O., Owen, B. D., Beresford-Jones, D., Moore, J. D., Kitchen, J. L. & Allaby, R. G. 2012. Archaeogenomic evidence of punctuated genome evolution in Gossypium. *Molecular Biology and Evolution,* **29**, (8): 2031-2038.
Panaro, N. J., Yuen, P. K., Sakazume, T., Fortina, P., Kricka, L. J. & Wilding, P. 2000. Evaluation of DNA fragment sizing and quantification by the Agilent 2100 Bioanalyzer. *Clinical Chemistry,* **46**, (11): 1851-1853.
Parmentier, I., Duminil, J., Kuzmina, M., Philippe, M., Thomas, D. W., Kenfack, D., Chuyong, G. B., Cruaud, C. & Hardy, O. J. 2013. How effective are DNA barcodes in the identification of African rainforest trees? *PLoS ONE,* **8**, (4): e54921.
Patterson, B. D., Willig, M. R. & Stevens, R. D. 2003. Trophic strategies, niche partitioning, and patterns of ecological organization*, University of Chicago Press.* 536-579.
Pearman, J. K., El-Sherbiny, M. M., Lanzén, A., Al-Aidaroos, A. M. & Irigoien, X. 2014. Zooplankton diversity across three Red Sea reefs using pyrosequencing. *Frontiers in Marine Science*.
Petney, T., Sithithaworn, P., Satrawaha, R., Grundy-Warr, C., Andrews, R., Wang, Y.-C. & Feng, C.-C. 2000. Fatal borreliosis in bat caused by relapsing fever spirochete, United Kingdom. *Health,* **31**: 225-37.
Petrov, B. P. 2006. Distribution and status of *Myotis bechsteinii* in Bulgaria (Chiroptera: Vespertilionidae). *Lynx,* **37**: 179-195.
Pianka, E. R. 1973. The structure of lizard communities. *Annual Review of Ecology and Systematics,* **4**: 53-74.
Pianka, E. R. 1981. Competition and niche theory. *Theoretical Ecology Principles and Applications*: 167-196.
Pineda-Munoz, S. & Alroy, J. 2014. Dietary characterization of terrestrial mammals. *Proceedings of the Royal Society of London B: Biological Sciences,* **281**, (1789): 20141173.
Poinar, H. 2002. DNA chemical instability. *eLS*: 1-9.
Poinar, H. N. & Cooper, A. 2000. Ancient DNA: do it right or not at all. *Science,* **5482**: 1139.
Polz, M. F. & Cavanaugh, C. M. 1998. Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology,* **64**, (10): 3724-3730.

Pompanon, F., Bonin, A., Bellemain, E. & Taberlet, P. 2005. Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics,* **6**, (11): 847-846.
Pompanon, F., Deagle, B. E., Symondson, W. O., Brown, D. S., Jarman, S. N. & Taberlet, P. 2012. Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology,* **21**, (8): 1931-1950.
Poulton, E. B. 1929. British insectivorous bats and their prey. *Proceedings of the Zoological Society of London,* **99**, (2): 277-303.
Prado, V., Castro, A., Oliveira, C., Souza, K. & Pena, S. 1997. Extraction of DNA from human skeletal remains: practical applications in forensic sciences. *Genetic Analysis: Biomolecular Engineering,* **14**, (2): 41-44.
Pruitt, K. D., Tatusova, T. & Maglott, D. R. 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research,* **35**, (suppl 1): D61-D65.
Puechmaille, S. J., Mathy, G. & Petit, E. J. 2007. Good DNA from bat droppings. *Acta Chiropterologica,* **9**, (1): 269-276.
Puechmaille, S. J., Wibbelt, G., Korn, V., Fuller, H., Forget, F., Mühldorfer, K., Kurth, A., Bogdanowicz, W., Borel, C. & Bosch, T. 2011. Pan-European distribution of white-nose syndrome fungus (*Geomyces destructans*) not associated with mass mortality. *PLoS ONE,* **6**, (4): e19167.
Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Bertoni, A., Swerdlow, H. P. & Gu, Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics,* **13**, (1): 1.
R Developement Core Team 2013. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
Rabinowitz, A. R. & Tuttle, M. D. 1982. A test of the validity of two currently used methods of determining bat prey preferences. *Acta Theriologica,* **27**, (13): 283-293.
Rajilić‐Stojanović, M., Smidt, H. & De Vos, W. M. 2007. Diversity of the human gastrointestinal tract microbiota revisited. *Environmental Microbiology,* **9**, (9): 2125-2136.
Ramulu, H. G., Groussin, M., Talla, E., Planel, R., Daubin, V. & Brochier-Armanet, C. 2014. Ribosomal proteins: Toward a next generation standard for prokaryotic systematics? *Molecular Phylogenetics and Evolution,* **75**: 103-117.
Ratnasingham, S. & Hebert, P. D. 2007. BOLD: The barcode of life data system (http://www.barcodinglife.org/). *Molecular Ecology Notes,* **7**, (3): 355-364.
Razgour, O., Clare, E. L., Zeale, M. R. K., Hanmer, J., Schnell, I. B., Rasmussen, M., Gilbert, T. P. & Jones, G. 2011. High‐throughput sequencing offers insight into mechanisms of resource partitioning in cryptic bat species. *Ecology and Evolution,* **1**, (4): 556-570.
Rebelo, H., Tarroso, P. & Jones, G. 2010. Predicted impact of climate change on European bats in relation to their biogeographic patterns. *Global Change Biology,* **16**, (2): 561-576.
Reyes, B., Wehdnesday, A., Rovira, H. G., Masangkay, J. S., Ramirez, T. J., Yoshikawa, Y. & Baticados, W. N. 2011. Polymerase chain reaction assay and conventional isolation of *Salmonella* spp. from Philippine bats. *Acta Scientiae Veterinariae,* **39**, (1).
Richardson, P. 2011. Bats*, Natural History Museum.* 1-128.
Riesenfeld, C. S., Schloss, P. D. & Handelsman, J. 2004. Metagenomics: genomic analysis of microbial communities. *Annual Review of Genetics,* **38**: 525-552.
Robert, V. A. & Casadevall, A. 2009. Vertebrate endothermy restricts most fungi as potential pathogens. *Journal Of Infectious Diseases,* **200**, (10): 1623-6.
Robinson, M. 1990. Prey selection by the brown long-eared bat (*Plecotus auritus*). *Myotis,* **28**: 5-18.
Robinson, M. & Stebbings, R. 1993. Food of the serotine bat, *Eptesicus serotinus*—is faecal analysis a valid qualitative and quantitative technique? *Journal of Zoology,* **231**, (2): 239-248.
Robinson, M. F. & Stebbings, R. E. 1997. Home range and habitat use by the serotine bat, *Eptesicus serotinus*, in England. *Journal of Zoology,* **243**, (1): 117-136.

Rohland, N. & Hofreiter, M. 2007. Comparison and optimization of ancient DNA extraction. *Biotechniques,* **42**, (3): 343.
Rohland, N. & Reich, D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research,* **22**, (5): 939-946.
Rostovskaya, M. S., Zhukova, D. V., Illarionova, A. E., Ustyugova, S. V., Borissenko, A. V. & Sviridov, A. V. 2000. Insect prey of the Long-eared bat *Plecotus auritus* (L.)(Chiroptera: Vespertilionidae) in Central Russia. *Russian Entomological Journal,* **9**: 185-189.
Roswag, A., Becker, N. I. & Encarnação, J. A. 2015. Importance of multi-dimensional analyses of resource partitioning in highly mobile species assemblages. *Population Ecology,* **57**, (4): 601-611.
Rozalska, B., Radzicki, G., Sadowska, B., Markowski, J. & Rudnicka, W. 1998. Aerobic microflora of *Myotis myotis* (Borkhausen, 1797) and *Barbastella barbastellus* (Schreber, 1774). *Bulletin of the Polish Academy of Sciences/Biological Sciences,* **46**: 59-67.
Ruano, G. & Kidd, K. K. 1992. Modeling of heteroduplex formation during PCR from mixtures of DNA templates. *Genome Research,* **2**, (2): 112-116.
Russ, J. & Racey, P. 2007. Species-specificity and individual variation in the song of male Nathusius' pipistrelles (*Pipistrellus nathusii*). *Behavioral Ecology and Sociobiology,* **61**, (5): 669-677.
Russo, D., Cistrone, L. & Jones, G. 2007. Emergence time in forest bats: the influence of canopy closure. *Acta Oecologica,* **31**, (1): 119-126.
Rust, M. & Reierson, D. 2007. Cockroaches: Integrated pest management for home gardeners and landscape professionals. *UC Statewide IPM Program, University of California, Davis*.
Rydell, J. 1989a. Feeding activity of the northern bat *Eptesicus nilssoni* during pregnancy and lactation. *Oecologia,* **80**, (4): 562-565.
Rydell, J. 1989b. Food habits of northern (*Eptesicus nilssoni*) and brown long‐eared (*Plecotus auritus*) bats in Sweden. *Ecography,* **12**, (1): 16-20.
Rydell, J., Bushby, A., Cosgrove, C. C. & Racey, P. A. 1994. Habitat use by bats along rivers in north east Scotland. *Folia Zoologica-Praha,* **43**: 417-417.
Rydell, J., Miller, L. & Jensen, M. 1999. Echolocation constraints of Daubenton's bat foraging over water. *Functional Ecology,* **13**, (2): 247-255.
Rydell, J., Natuschke, G., Theiler, A. & Zingg, P. E. 1996. Food habits of the barbastelle bat *Barbastella barbastellus*. *Ecography,* **19**, (1): 62-66.
Safi, K. & Kerth, G. 2004. A comparative analysis of specialization and extinction risk in temperate‐zone bats. *Conservation Biology,* **18**, (5): 1293-1303.
Salter, S., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. & Walker, A. W. 2014a. Reagent contamination can critically impact sequence-based microbiome analyses. *bioRxiv*: 007187.
Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J. & Walker, A. W. 2014b. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology,* **12**, (1): 87.
Salyers, A., Vercellotti, J., West, S. & Wilkins, T. 1977. Fermentation of mucin and plant polysaccharides by strains of Bacteroides from the human colon. *Applied and Environmental Microbiology,* **33**, (2): 319-322.
Sandström, J. P., Russell, J. A., White, J. P. & Moran, N. A. 2001. Independent origins and horizontal transfer of bacterial symbionts of aphids. *Molecular Ecology,* **10**, (1): 217-228.
Sattler, T., Bontadina, F., Hirzel, A. H. & Arlettaz, R. 2007. Ecological niche modelling of two cryptic bat species calls for a reassessment of their conservation status. *Journal of Applied Ecology,* **44**, (6): 1188-1199.
Sboner, A., Mu, X. J., Greenbaum, D., Auerbach, R. K. & Gerstein, M. B. 2011. The real cost of sequencing: higher than you think! *Genome Biology,* **12**, (8): 125.
Schmieder, D. A., Benítez, H. A., Borissov, I. M. & Fruciano, C. 2015. Bat species comparisons based on external morphology: a test of traditional versus geometric morphometric approaches. *PLoS ONE,* **10**, (5).

Schnitzler, H.-U., Moss, C. F. & Denzinger, A. 2003. From spatial orientation to food acquisition in echolocating bats. *Trends in Ecology & Evolution,* **18**, (8): 386-394.
Schoener, T. W. 1974. Resource partitioning in ecological communities. *Science,* **185**, (4145): 27-39.
Scholz, M. B., Lo, C.-C. & Chain, P. S. 2012. Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. *Current Opinion in Biotechnology,* **23**, (1): 9-15.
Sebek, Z. 1974. Blood parasites of small wild mammals in Czechoslovakia. *Folia Parasitologica,* **22**, (1): 11-20.
Seim, I., Fang, X., Xiong, Z., Lobanov, A. V., Huang, Z., Ma, S., Feng, Y., Turanov, A. A., Zhu, Y. & Lenz, T. L. 2013. Genome analysis reveals insights into physiology and longevity of the Brandt's bat *Myotis brandtii*. *Nature Communications,* **4**, (2212): 1-8.
Shahack-Gross, R., Berna, F., Karkanas, P. & Weiner, S. 2004. Bat guano and preservation of archaeological remains in cave sites. *Journal of Archaeological Science,* **31**, (9): 1259-1272.
Shannon, C. E. & Weaver, W. 1948. A mathematical theory of communication. *ACM SIGMOBILE,* **5**, (1): 3-55.
Shendure, J. & Ji, H. 2008. Next-generation DNA sequencing. *Nature Biotechnology,* **26**, (10): 1135-1145.
Sherley, M., Gordon, D. M. & Collignon, P. J. 2000. Variations in antibiotic resistance profile in Enterobacteriaceae isolated from wild Australian mammals. *Environmental Microbiology,* **2**, (6): 620-631.
Shiel, C., Mcaney, C. & Fairley, J. 1991. Analysis of the diet of Natterer's bat *Myotis nattereri* and the common long‐eared bat *Plecotus auritus* in the west of Ireland. *Journal of Zoology,* **223**, (2): 299-305.
Shiel, C., Mcaney, C. & Sullivan, C. 1997. Identification of arthropod fragments in bat droppings. *Mammal Society*.
Shiel, C. B., Duverge, P. L., Smiddy, P. & Fairley, J. S. 1998. Analysis of the diet of Leisler's bat (*Nyctalus leisleri*) in Ireland with some comparative analyses from England and Germany. *Journal of Zoology,* **246**, (04): 417-425.
Siemers, B. M. & Schnitzler, H.-U. 2000. Natterer's bat (*Myotis nattereri,* Kuhl, 1818) hawks for prey close to vegetation using echolocation signals of very broad bandwidth. *Behavioral Ecology and Sociobiology,* **47**, (6): 400-412.
Siemers, B. M. & Swift, S. M. 2006. Differences in sensory ecology contribute to resource partitioning in the bats *Myotis bechstein*ii and *Myotis nattereri* (Chiroptera: Vespertilionidae). *Behavioral Ecology and Sociobiology,* **59**, (3): 373-380.
Sierro, A. & Arlettaz, R. 1997. Barbastelle bats (*Barbastella spp.*) specialize in the predation of moths: implications for foraging tactics and conservation. *Acta Oecologica,* **18**, (2): 91-106.
Simmons, N. B. 1993. Morphology, function, and phylogenetic significance of pubic nipples in bats (Mammalia, Chiroptera). *American Museum Novitates,* **3077**: 1-37.
Simmons, N. B. & Geisler, J. H. 1998. Phylogenetic relationships of Icaronycteris, Archaeonycteris, Hassianycteris, and Palaeochiropteryx to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bulletin of the AMNH,* **235**: 2-180.
Sipos, R., Székely, A. J., Palatinszky, M., Révész, S., Márialigeti, K. & Nikolausz, M. 2007. Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiology Ecology,* **60**, (2): 341-350.
Skerra, A. 1992. Phosphorothioate primers improve the amplification of DNA sequences by DNA polymerases with proofreading activity. *Nucleic Acids Research,* **20**, (14): 3551-3554.
Smirnov, D. & Vekhnik, V. 2014. Ecology of nutrition and differentiation of the trophic niches of bats (Chiroptera: Vespertilionidae) in floodplain ecosystems of the Samara Bend. *Biology Bulletin,* **41**, (1): 60-70.

Smith, C. I., Chamberlain, A. T., Riley, M. S., Stringer, C. & Collins, M. J. 2003. The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution,* **45**, (3): 203-217.
Smith, O., Momber, G., Bates, R., Garwood, P., Fitch, S., Pallen, M., Gaffney, V. & Allaby, R. G. 2015. Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000 years ago. *Science,* **347**, (6225): 998-1001.
Sologor, E. A. 1980. Kizuchenyu pitaniya *Vespertilio serotinus*. *Voprosy Teriologii. Rukokrylye*: 188-190.
Sommer, F., Ståhlman, M., Ilkayeva, O., Arnemo, J. M., Kindberg, J., Josefsson, J., Newgard, C. B., Fröbert, O. & Bäckhed, F. 2016. The gut microbiota modulates energy metabolism in the hibernating brown bear *Ursus arctos*. *Cell Reports,* **14**, (7): 1655-1661.
Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C.-H., Westover, B. P., Weatherford, J., Buhler, J. D. & Gordon, J. I. 2005. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science,* **307**, (5717): 1955-1959.
Sparks, T., Roy, D. & Dennis, R. 2005. The influence of temperature on migration of Lepidoptera into Britain. *Global Change Biology,* **11**, (3): 507-514.
Sparks, T. H., Dennis, R. L., Croxton, P. J. & Cade, M. 2007. Increased migration of Lepidoptera linked to climate change. *European Journal of Entomology,* **104**, (1): 139.
Speakman, J. R. & Rowland, A. 1999. Preparing for inactivity: how insectivorous bats deposit a fat store for hibernation. *Proceedings of the Nutrition Society,* **58**, (01): 123-131.
Srivathsan, A., Sha, J., Vogler, A. P. & Meier, R. 2015. Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf‐feeding monkey (*Pygathrix nemaeus*). *Molecular Ecology Resources,* **15**, (2): 250-261.
Stackebrandt, E., Rainey, F. A. & Ward-Rainey, N. L. 1997. Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *International Journal of Systematic and Evolutionary Microbiology,* **47**, (2): 479-491.
Steane, A. M. 1996. Error correcting codes in quantum theory. *Physical Review Letters,* **77**, (5): 793-797.
Stebbings, R. E. 1986. Which bat Is It? A guide to bat identification in Great Britain and Ireland*, The Mammal Society and the Vincent Wildlife Trust.* 32-34.
Stork, N. E., Mcbroom, J., Gely, C. & Hamilton, A. J. 2015. New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods. *Proceedings of the National Academy of Sciences,* **112**, (24): 7519-7523.
Sullivan, C., Shiel, C., Mcaney, C. & Fairley, J. 1993. Analysis of the diets of Leisler's *Nyctalus leisleri,* Daubenton's *Myotis daubentoni* and Pipistrelle *Pipistrellus pipistrellus* bats in Ireland. *Journal of Zoology,* **231**, (4): 656-663.
Swift, S. & Racey, P. 1983. Resource partitioning in two species of vespertilionid bats (Chiroptera) occupying the same roost. *Journal of Zoology,* **200**, (2): 249-259.
Swift, S. & Racey, P. 2002. Gleaning as a foraging strategy in Natterer's bat *Myotis nattereri*. *Behavioral Ecology and Sociobiology,* **52**, (5): 408-416.
Swift, S., Racey, P. & Avery, M. 1985. Feeding ecology of *Pipistrellus pipistrellus* (Chiroptera: Vespertilionidae) during pregnancy and lactation. II. Diet. *The Journal of Animal Ecology,* **54**: 217-225.
Taake, K. 1993. Zur Nahrungsökologie waldbewohnender Fledermäuse (Chiroptera, Vespertilionidae)—ein Nachtrag. *Myotis,* **31**: 163-165.
Taberlet, P., Waits, L. P. & Luikart, G. 1999. Noninvasive genetic sampling: look before you leap. *Trends in Ecology & Evolution,* **14**, (8): 323-327.
Tayyab, S. 1992. Essential molecular biology: A practical approach volume I: Edited by TA Brown. The Practical Approach Series, IRL Press, Oxford. 1991: ISBN 0‐19‐963110‐7. *Biochemical Education,* **20**, (3): 155-156.
Teeling, E. C. 2009. Hear, hear: the convergent evolution of echolocation in bats? *Trends in Ecology & Evolution,* **24**, (7): 351-354.
Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O'brien, S. J. & Murphy, W. J. 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science,* **307**, (5709): 580-584.

Ter Hofstede, H. M., Goerlitz, H. R., Ratcliffe, J. M., Holderied, M. W. & Surlykke, A. 2013. The simple ears of noctuoid moths are tuned to the calls of their sympatric bat community. *The Journal of Experimental Biology,* **216**, (21): 3954-3962.
Thompson, J. D., Gibson, T. & Higgins, D. G. 2002. Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*: 2.3. 1-2.3. 22.
Thompson, M. 1982. A common long-eared bat *Plecotus auritus*: moth predator-prey relationship. *Naturalist,* **107**: 87-97.
Timm, R. M. & Genoways, H. H. 2004. The Florida bonneted bat, *Eumops floridanus* (Chiroptera: Molossidae): distribution, morphometrics, systematics, and ecology. *Journal of Mammalogy,* **85**, (5): 852-865.
Tollit, D. J., Schulze, A. D., Trites, A. W., Olesiuk, P. F., Crockford, S. J., Gelatt, T. S., Ream, R. R. & Miller, K. M. 2009. Development and application of DNA techniques for validating and improving pinniped diet estimates. *Ecological Applications,* **19**, (4): 889-905.
Tringe, S. G., Von Mering, C., Kobayashi, A., Salamov, A. A., Chen, K., Chang, H. W., Podar, M., Short, J. M., Mathur, E. J. & Detter, J. C. 2005. Comparative metagenomics of microbial communities. *Science,* **308**, (5721): 554-557.
Troyer, K. 1984. Diet selection and digestion in *Iguana iguana*: the importance of age and nutrient requirements. *Oecologia,* **61**, (2): 201-207.
Turner, J. R., Gatehouse, C. M. & Corey, C. A. 1987. Does solar energy control organic diversity? Butterflies, moths and the British climate. *Oikos*: 195-205.
Van Den Bussche, R. A. & Hoofer, S. R. 2004. Phylogenetic relationships among recent chiropteran families and the importance of choosing appropriate out-group taxa. *Journal of Mammalogy,* **85**, (2): 321-330.
Van Der Wilk, F., Dullemans, A. M., Verbeek, M. & Van Den Heuvel, J. F. 1999. Isolation and characterization of *AP*SE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrthosiphon pisum*. *Virology,* **262**, (1): 104-113.
Vaughan, N. 1997. The diets of British bats (Chiroptera). *Mammal Review,* **27**, (2): 77-94.
Verant, M. L., Boyles, J. G., Waldrep Jr, W., Wibbelt, G. & Blehert, D. S. 2012. Temperature-dependent growth of *Geomyces destructans*, the fungus that causes bat white-nose syndrome. *PLoS ONE,* **7**, (9): e46280.
Vesterinen, E. J., Lilley, T., Laine, V. N. & Wahlberg, N. 2013. Next generation sequencing of fecal DNA reveals the dietary diversity of the widespread insectivorous predator Daubenton's Bat (*Myotis daubentonii*) in Southwestern Finland. *PLoS ONE,* **8**, (11): e82168.
Von Helversen, O., Heller, K.-G., Mayer, F., Nemeth, A., Volleth, M. & Gombkötö, P. 2001. Cryptic mammalian species: a new species of whiskered bat (*Myotis alcathoe n. sp.*) in Europe. *Naturwissenschaften,* **88**, (5): 217-223.
Waage, J. K. 1979. The evolution of insect/vertebrate associations. *Biological Journal of the Linnean Society,* **12**, (3): 187-224.
Walters, C. L., Collen, A., Lucas, T., Mroz, K., Sayer, C. A. & Jones, K. E. 2013. Challenges of using bioacoustics to globally monitor bats*, Bat Evolution, Ecology, and Conservation, Springer.* 479-499.
Wang Jing, W. X., Jiang Tinglei,Wang Lei,Lu Guanjun,You Yuyan,Liu Ying,Li Dan,Feng Jiang 2010. Relationships between foraging activity of greater horseshoe bat (*Rhinolophus ferrumequinum*) and prey resources. *Acta Theriologica Sinica,* **30**, (2): 157-162.
Wang, L.-F. & Cowled, C. 2015. Bats and viruses: a new frontier of emerging infectious diseases*, John Wiley & Sons.*
Wang, Y. & Qian, P.-Y. 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE,* **4**, (10): e7401.
Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* **360**, (1462): 1847-1857.
Waters, D., Jones, G. & Furlong, M. 1999. Foraging ecology of Leisler's bat (*Nyctalus leisleri*) at two sites in southern Britain. *Journal of Zoology,* **249**, (2): 173-180.

Waters, D. A., Rydell, J. & Jones, G. 1995. Echolocation call design and limits on prey size: a case study using the aerial-hawking bat *Nyctalus leisleri*. *Behavioral Ecology and Sociobiology,* **37**, (5): 321-328.
Wetterstrand, K. 2015. *DNA sequencing costs: data from the NHGRI Genome sequencing program (GSP)* [Online]. Available: http://www.genome.gov/sequencingcosts [Accessed 6 January 2016].
Whitaker, J. O., Dannelly, H. K. & Prentice, D. A. 2004. Chitinase in insectivorous bats. *Journal of Mammalogy,* **85**, (1): 15-18.
Whitaker Jr, J. O. & Karatas, A. 2009. Food and feeding habits of some bats from Turkey. *Acta Chiropterologica,* **11**, (1): 393-403.
Wiedbrauk, D. L., Werner, J. C. & Drevon, A. M. 1995. Inhibition of PCR by aqueous and vitreous fluids. *Journal of Clinical Microbiology,* **33**, (10): 2643-2646.
Williams, C., Salter, L. & Jones, G. 2010. The winter diet of the lesser horseshoe bat (*Rhinolophus hipposideros*) in Britain and Ireland. *Hystrix, the Italian Journal of Mammalogy,* **22**, (1).
Willis, C. & Brigham, R. 2003. Defining torpor in free-ranging bats: experimental evaluation of external temperature-sensitive radiotransmitters and the concept of active temperature. *Journal of Comparative Physiology B,* **173**, (5): 379-389.
Wilson, D. E. & Reeder, D. M. 2005. Mammal species of the world: a taxonomic and geographic reference*, JHU Press.* 12.
Wolz, I. 1993. Das beutespektrum der bechsteinfledermaus *Myotis bechsteini* (Kuhl, 1818) ermittelt aus kotanalysen. *Myotis,* **31**: 27-68.
Wong, S., Lau, S., Woo, P. & Yuen, K. Y. 2007. Bats as a continuing source of emerging infections in humans. *Reviews in Medical Virology,* **17**, (2): 67-91.
Yager, D. D. 2012. Predator detection and evasion by flying insects. *Current Opinion in Neurobiology,* **22**, (2): 201-207.
Yoon, K. B., Kim, J. Y., Cho, J. Y. & Park, Y. C. 2011. The complete mitochondrial genome of the greater horseshoe bat subspecies, *Rhinolophus ferrumequinum korai* (Chiroptera: Rhinolophidae). *Mitochondrial DNA,* **22**, (4): 102-104.
Yoon, K. B., Kim, J. Y., Kim, H. R., Cho, J. Y. & Park, Y. C. 2013. The complete mitochondrial DNA genome of a greater horseshoe bat subspecies, *Rhinolophus ferrumequinum quelpartis* (Chiroptera: Rhinolophidae). *Mitochondrial DNA,* **24**, (1): 22-24.
Young, R. A. 1976. Fat, energy and mammalian survival. *American Zoologist,* **16**, (4): 699-710.
Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., Lees, D. C. & Jones, G. 2011. Taxon‐specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources,* **11**, (2): 236-244.
Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. 2000. A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology,* **7**, (1-2): 203-214.
Zukal, J. & Gajdošík, M. 2012. Diet of *Eptesicus serotinus* in an agricultural landscape. *Vespertilio,* **16**: 357-363.